

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:19:54 ; Search time 26 Seconds
(without alignments)
20.500 Million cell updates/sec

Title: US-10-033-526-2
Database: 1 AAPM 4
Sequence: 1 AAPM 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|------------------------|
| 1 | 20 | 100.0 | 4 | 5 | AAP40460 | N-terminal polypeptide |
| 2 | 20 | 100.0 | 4 | 13 | AAR22056 | Chromogenic substr |
| 3 | 20 | 100.0 | 4 | 13 | AAR26241 | Alpha-keto peptide |
| 4 | 20 | 100.0 | 4 | 15 | AAR46222 | Serine protease in |
| 5 | 20 | 100.0 | 4 | 18 | AAW52600 | Serine protease-in |
| 6 | 20 | 100.0 | 4 | 18 | AAW52609 | Serine protease-in |
| 7 | 20 | 100.0 | 4 | 18 | AAW29407 | Cathepsin G inhibi |
| 8 | 20 | 100.0 | 4 | 18 | AAW26078 | Substrate #1 for s |
| 9 | 20 | 100.0 | 4 | 19 | AAW51453 | Indicator for dete |
| 10 | 20 | 100.0 | 4 | 19 | AAW51449 | Indicator for dete |

| | | | | | | |
|----|----|-------|----|----|----------|--------------------|
| 11 | 20 | 100.0 | 4 | 19 | AAW51611 | Peptide conjugated |
| 12 | 20 | 100.0 | 4 | 20 | AAV31239 | Alpha-aminoalkyl p |
| 13 | 20 | 100.0 | 4 | 20 | AAW82444 | Seq ID 13 from US5 |
| 14 | 20 | 100.0 | 4 | 22 | AAW97634 | Alpha-ketoamide ca |
| 15 | 20 | 100.0 | 4 | 22 | AAW82199 | Cathepsin G recogn |
| 16 | 20 | 100.0 | 4 | 22 | AAW62788 | Peptide-lipid conf |
| 17 | 20 | 100.0 | 4 | 23 | ABB83504 | Enzyme cleavable p |
| 18 | 20 | 100.0 | 4 | 23 | ABB83518 | Enzyme cleavable p |
| 19 | 20 | 100.0 | 4 | 23 | AAO18044 | C-terminal truncat |
| 20 | 20 | 100.0 | 5 | 23 | ABB83531 | Enzyme cleavable p |
| 21 | 20 | 100.0 | 5 | 23 | ABB83532 | Enzyme cleavable p |
| 22 | 20 | 100.0 | 9 | 21 | AAW6270 | Human proBDNF sign |
| 23 | 20 | 100.0 | 14 | 21 | AAW10855 | pSectag-VEGF prote |
| 24 | 20 | 100.0 | 15 | 22 | AAW98208 | Human p24 protein- |
| 25 | 20 | 100.0 | 17 | 14 | AAW42608 | Encoded by human V |
| 26 | 20 | 100.0 | 17 | 17 | AAW94034 | VEGF exon II. Hom |
| 27 | 20 | 100.0 | 19 | 22 | AAW00655 | Human targeting pe |
| 28 | 20 | 100.0 | 25 | 23 | ABG62530 | Eubacterial MutSI |
| 29 | 20 | 100.0 | 31 | 22 | AAW19115 | Peptide #5549 enco |
| 30 | 20 | 100.0 | 31 | 22 | AAW31764 | Peptide #5801 enco |
| 31 | 20 | 100.0 | 31 | 23 | ABG41275 | Human secreted pro |
| 32 | 20 | 100.0 | 42 | 21 | AAW39021 | Human secreted pro |
| 33 | 20 | 100.0 | 43 | 22 | ABG27642 | Novel human diagno |
| 34 | 20 | 100.0 | 50 | 22 | AAW64697 | Propionibacterium |
| 35 | 20 | 100.0 | 50 | 22 | AAW65105 | Propionibacterium |
| 36 | 20 | 100.0 | 51 | 21 | AAW65229 | Human 5' EST relat |
| 37 | 20 | 100.0 | 53 | 22 | ABG00125 | Novel human diagno |
| 38 | 20 | 100.0 | 56 | 22 | AAW64950 | Propionibacterium |
| 39 | 20 | 100.0 | 59 | 22 | AAW67577 | Propionibacterium |
| 40 | 20 | 100.0 | 60 | 23 | ABP10723 | Human OREF protein |
| 41 | 20 | 100.0 | 61 | 23 | ABP08952 | Human OREF protein |
| 42 | 20 | 100.0 | 66 | 22 | AAO12340 | Human polypeptide |
| 43 | 20 | 100.0 | 67 | 21 | AAW34555 | Arabidopsis thalia |
| 44 | 20 | 100.0 | 68 | 22 | AAW46479 | Propionibacterium |
| 45 | 20 | 100.0 | 68 | 22 | AAW64090 | Propionibacterium |

ALIGNMENTS

| | |
|--|--|
| RESULT 1 | |
| AAP40460 | |
| ID AAP40460 standard; Protein: 4 AA. | |
| XX | |
| AC AAP40460; | |
| XX | |
| DT 27-NOV-1991 (first entry) | |
| XX | |
| DE N-terminal polyhydroxyalkanoyl peptide. | |
| XX | |
| KW Polyhydroxyalkanoyl peptide; protease; peptidase; prodrug. | |
| XX | |
| PN EP126685-A. | |
| XX | |
| PD 28-NOV-1985. | |
| XX | |
| PF 15-MAY-1984; 84EP-0400984. | |
| XX | |
| PR 16-MAY-1983; 83FR-0308051. | |
| XX | |
| PA (CNRS) CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE. | |
| XX | |
| PI Monsigny M, Mayer R; | |
| XX | |
| DR WPI; 1984-296065/48. | |
| XX | |
| PT N-terminal polyhydroxyalkanoyl peptide and aminoacid derivs. - opt. | |
| PT having C-terminal active gps., e.g. chloroquine, are water-soluble | |
| PT antitumour or antiparasitic cpds. and protease targets. | |
| XX | |
| PS Claim 11; page 15; 20pp; french. | |
| XX | |
| CC The peptide is polyhydroxyalkanoyl-L-Ala-L-Ala-L-Pro-L-Met-X. X is | |

CC R2-NH- derived from an aromatic amine, eg p-nitro-aniline, alpha-
CC or beta-naphthylamine, or it is R3NH derived from an amino drug, eg
CC daunorubicin or chloroquine, or it is a gp. which may confer
CC activity, eh H, OH, CH2Cl. When X is the residue of a drug, the
CC cpd. is a prodrug in which the peptide is the substrate for a
CC specific protease secreted by the target cell for the active drug,
CC eg tumour cells or microorganism pathogens, eg Plasmodium
CC falciparum. Due to the polyhydroxy gp., the cpds. are very soluble
CC in water or buffer soln. and are able to bu used at high concn. The
CC cpds. also allow the detection of proteases and peptidases and
CC allow easy determination of the best substrate for a particular
CC protease.
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
| | | |
Db 1 AAPM 4

RESULT 2
AAR22056
ID AAR22056 standard; Peptide; 4 AA.
XX
AC AAR22056;

DT 06-JUL-1992 (first entry)

XX Chromogenic substrate for cysteine proteases.

DE Beta amyloid; protein precursor; Alzheimers disease; assay.

KW Synthetic.

OS
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acylation by MeoSuc"
FT Modified-site 4 /note= "C-terminal modification by PMA
FT (para nitro amide)"
XX

XX WO9203542-A.

XX 05-MAR-1992.

XX 19-AUG-1991; 91WO-UO05932.

XX 17-AUG-1990; 90US-0568806.

XX (UYBO-) BOSTON UNIV.

XX Abraham CR;

XX WPI; 1992-096886/12.

XX
PT Treatment and diagnosis of Alzheimer's disease - by reducing
PT beta-protein precursor proteolysis near beta-protein N-terminus
PT by administering proteolysis inhibitor
XX
PS Disclosure; Page 10; 29pp; English.

XX The chromogenic peptide was used to assay the activity of purified
CC cysteine protease from Alzheimers disease patients. The protease
CC was incubated with the chromogenic substrate and changes in
CC absorbance followed at 410 nm in a Titertek Multiskan ELISA reader.
CC See also AAR22054,5.

XX Sequence 4 AA;

Query Match 100.0%; Score 20; DB 13; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
| | | |
Db 1 AAPM 4

RESULT 3
AAR26241
ID AAR26241 standard; peptide; 4 AA.
XX
AC AAR26241;

DT 27-JAN-1993 (first entry)

XX Alpha-keto peptide protease inhibitors.

XX serine protease inhibitor; cysteine protease inhibitor; chymase;
KW chymotrypsin; elastase; trypsin; blood coagulation enzyme;
KW antiinflammatory; anticoagulant; neuroprotectant; emphysema; ARDS;
KW arthritis; ischaemia; stroke; Alzheimer's disease; tumour;
KW metastases; bone resorption.

XX Synthetic.

OS
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "MeO-Suc-Ala"
FT Modified-site 4 /note= "Met-COQR, where QR denotes residue of a
FT keto-acid (OH), keto-amide (NH2) or keto-ester"
XX

XX WO9212140-A.

XX 23-JUL-1992.

XX 27-DEC-1991; 91WO-US09801.

XX 28-DEC-1990; 90US-0635287.

XX (GEOR-) GEORGIA TECH RES CORP.

XX Powers JC;

XX WPI; 1992-268589/32.

XX
PT New alpha-keto:amide, alpha-keto:acid and alpha-keto:ester
PT derivs. - inhibit serine and cysteine protease(s) and are used
PT for treating and preventing inflammation, emphysema, adult
PT respiratory distress syndrome, etc.

XX Disclosure; Page 31; 88pp; English.

XX The peptide is one example of a highly generic group of amino acids
CC and di-, tri- and tetra-peptides having alpha-keto functionality at
CC the C-terminal, the compounds optionally being in amide or ester form.
CC These compounds are selective or general inhibitors of serine and
CC cysteine proteases, e.g. chymases, chymotrypsin-like enzymes,
CC elastases, trypsin, and blood coagulation enzymes. They are useful
CC as antiinflammatory agents, anticoagulants, neuroprotectants and for
CC treatment of neurodegeneration in cases of e.g. emphysema, ARDS,
CC arthritis, ischaemia, stroke, Alzheimer's disease, tumour metastases
CC and bone resorption. They are also useful in-vitro for protecting
CC sensitive materials against proteolysis.

XX The present compound is specifically a cathepsin G inhibitor.

XX Sequence 4 AA;

Query Match 100.0%; Score 20; DB 13; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
1111
Db 1 AAPM 4

RESULT 4
AAR46222 standard; peptide; 4 AA.
XX
AC AAR46222;
XX
DT 04-AUG-1994 (first entry)
XX
DE Serine protease inhibitor tetrapeptide.
XX
KW Prevention; schistosomiasis; parasite; infection; prevention;
KW parasitic penetration; skin; cercariae; anti-penetrant.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Modified-site 1 /note= "Bg(peptide blocking gp.) attached"
FT Modified-site 4 /note= "PI(protease inhibitor), other than
FT chloromethyl ketone, attached"
XX
PN US5284829-A.
XX
PD 08-FEB-1994.
XX
PF 26-NOV-1991; 91US-0798565.
XX
PR 26-NOV-1991; 91US-0798565.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Cohen FE, Mckerrow JH;
XX
DR WPI; 1994-056364/07.
XX
PT Synthetic tetra:peptide(s) with an N-terminal blocking gp. and
PT C-terminal enzyme inhibitor - can be formulated into soaps and
PT sprays and used to prevent schistosomal skin penetration
XX
PS Disclosure; Page 7; 35pp; English.
XX
CC The sequence is that of a synthetic tetrapeptide serine protease
CC inhibitor which can be used to prevent schistosome parasite
CC infection. It may be used in a formulation as a soap, lotion,
CC cream, spray, etc. to stop parasitic penetration of the skin.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 15; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
1111
Db 1 AAPM 4

RESULT 5
AAW52600 standard; peptide; 4 AA.
XX
AC AAW52600;
XX
DT 22-JUN-1998 (first entry)
XX
DE Serine protease-inhibiting peptide with C-terminal phosphonate residue.
XX

KW Selective; serine protease inhibitor; trypsin; elastase; chymotrypsin;
KW antiinflammatory; anticoagulant; antitumour; cathepsin G.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Modified-site 1 /note= "Meo-Suc-Ala"
FT Modified-site 4 /note= "Met(P)-(Oph)2; where Met(P) represents a
FT methionine analogue in which the carboxyl group
FT -COOH has been replaced by a phosphonate group
FT -P(=O)(OH)(OH); and (Oph)2 indicates that the
FT phosphonate has been diphenyl esterified"
XX
PN US5686419-A.
XX
PD 11-NOV-1997.
XX
PE 21-JAN-1994; 94US-0184286.
XX
PR 21-JAN-1994; 94US-0184286.
XX
PA (GEOR-) GEORGIA TECH RES CORP.
XX
PI Boduszek B, Oleksyszyn J, Powers JC;
XX
DR WPI; 1997-558177/51.
XX
PT New peptidyl derivatives of alpha-aminoalkyl phosphonic acid
PT di:ester - are serine protease inhibitors, useful e.g. for reducing
PT blood coagulation, controlling tumour invasion or treating
PT inflammation
XX
PS Example 7; Column 20; 16pp; English.
XX
CC The patent discloses new peptidyl derivatives of diesters of alpha-
CC aminoalkylphosphonic acids having basic substituents, of formula
CC X-AA4-AA3-AA2-NH-CHR-P(=O)(OZ)(OZ1), in which: R = B-substituted
CC phenyl, B-substituted benzyl or B-substituted 1-6C alkyl; B = amidino,
CC guanidino, isothioureido or amino; Z and Z1 = 1-6C perfluoroalkyl or
CC optionally substituted phenyl; AA2 = an L- or D-amino acid residue;
CC AA3 = a single bond or an L- or D-amino acid residue; AA4 = a single
CC bond or an L- or D-amino acid residue; X = H, NH2CO, NH2CS, NH2SO2,
CC YNHCO, YNHCS, YNHSO2, YCS, YSO2, YOCO, YOCS or YCO; Y = optionally
CC substituted 1-6C alkyl, 1-6C fluoroalkyl, 9-fluorenylmethyl, phenyl,
CC naphthyl, or 1-6C alkyl with 1-2, optionally substituted, attached
CC phenyl groups. The new compounds are inhibitors of trypsin, elastase,
CC chymotrypsin and other serine proteases; and their inhibitory activity
CC is selective depending on the identity of the alpha-aminoalkylphosphonic
CC acid ester residue. They can be used as antiinflammatory agents,
CC anticoagulants and antitumour agents. The present sequence is an
CC example of a compound which can specifically inhibit the activity of
CC Cathepsin G.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 18; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
1111
Db 1 AAPM 4

RESULT 6
AAW52609 standard; peptide; 4 AA.
XX
AC AAW52609;
XX
DT 22-JUN-1998 (first entry)
XX

XX Serine protease-inhibiting peptide with C-terminal phosphonate residue.
DE
XX
KW Selective; serine protease inhibitor; trypsin; elastase; chymotrypsin;
XX antiinflammatory; anticoagulant; antitumour.
OS
XX Synthetic.
FH
FT Key Location/Qualifiers
FT Modified-site 1
FT Modified-site 4 /note= "MeO-Suc-Ala"
FT
FT /note= "Met(O)(P)-(OPh)2; where Met(O)(P) represents a
FT methionine sulphoxide analogue in which the carboxyl
FT group -COOH has been replaced by a phosphonate group
FT -P(=O)(OH)(OH); and (OPh)2 indicates that the
FT phosphonate has been diphenyl esterified"
XX
PN US5686419-A.
XX
PD 11-NOV-1997.
XX
PF 21-JAN-1994; 94US-0184286.
XX
PR 21-JAN-1994; 94US-0184286.
XX
PA (GEOR-) GEORGIA TECH RES CORP.
PI
PI Boduszek B, Oleksyszyn J, Powers JC;
XX
XX WPI; 1997-558177/51.
DR
XX
XX New peptidyl derivatives of alpha-aminoalkyl phosphonic acid
PT di:ester - are serine protease inhibitors, useful e.g. for reducing
PT blood coagulation, controlling tumour invasion or treating
PT inflammation
XX
XX
PS Disclosure; Column 11; 16pp; English.
XX
XX The patent discloses new peptidyl derivatives of diesters of alpha-
CC aminoalkylphosphonic acids having basic substituents, of formula
CC X-AA4-AA3-AA2-NH-CHR-P(=O)(OZ)(OZ1), in which: R = B-substituted
CC phenyl, B-substituted benzyl or B-substituted 1-6C alkyl; B = amidino,
CC guanidino, isothioureido or amino; Z and Z1 = 1-6C perfluoroalkyl or
CC optionally substituted phenyl; AA2 = an L- or D-amino acid residue;
CC AA3 = a single bond or an L- or D-amino acid residue; AA4 = a single
CC bond or an L- or D-amino acid residue; X = H, NH2CO, NH2CS, NH2SO2,
CC YNHCO, YNHCS, YNHCO2, YCS, YSO2, YOCO, YOCS or YCO; Y = optionally
CC substituted 1-6C alkyl, 1-6C fluoroalkyl, 9-fluorenylmethyl, phenyl,
CC naphthyl, or 1-6C alkyl with 1-2, optionally substituted, attached
CC phenyl groups. The new compounds are inhibitors of trypsin, elastase,
CC chymotrypsin and other serine proteases; and their inhibitory activity
CC is selective depending on the identity of the alpha-aminoalkylphosphonic
CC acid ester residue. They can be used as antiinflammatory agents,
CC anticoagulants and antitumour agents. The present sequence is a
CC specific example of the new compounds.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 18; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 1 AAPM 4

RESULT 7
AAW29407
ID AAW29407 standard; peptide; 4 AA.
XX
AC AAW29407;

XX
DT 23-FEB-1998 (first entry)
XX
DE Cathepsin G inhibitor peptide ketoamide derivative.
XX
KW Peptide ketoamide derivative; protease inhibitor; elastase inhibitor;
KW Alzheimer's disease; coagulation disorder; serine protease;
KW cysteine protease; calpain; cathepsin G; neurodegenerative disease;
KW ischaemia; stroke; trypsin inhibitor; chymase inhibitor;
KW tissue damage; thrombosis; blistering; anticoagulant.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site 4 /note= "MeO-succinyl-Ala"
FT
FT /note= "Met-NR3R4, where R3 and R4 are selected
FT independently from the group consisting of H,
FT 1-20C alkyl, 3-20C cyclised alkyl, 1-20C
FT alkyl with a phenyl group (optionally mono-,
FT di- or tri-substituted with K) attached to the
FT 1-20C alkyl, 3-20C cyclised alkyl with an
FT attached phenyl group substituted with K, 1-20
FT alkyl with a morpholine, a pyrrolidine or a
FT piperidine ring attached through nitrogen to the
FT alkyl, 1-20C alkyl with an OH group attached to
FT the alkyl, 1-10C with an attached pyridyl
FT group or cyclohexyl group, -NCH2CH2-
FT (4-hydroxyphenyl) - and -NCH2CH2-(3-indolyl)"
XX
XX US5610297-A.
PN
XX
XX 11-MAR-1997.
PD
XX
PF 06-OCT-1995; 95US-0815073.
XX
XX 27-DEC-1991; 91US-0815073.
PR
PR 09-SEP-1993; 93US-0118997.
PR 20-MAY-1994; 94US-0247081.
XX
XX (GEOR-) GEORGIA TECH RES CORP.
PA
PI Powers JC;
PI WPI; 1997-178454/16.
DR
XX
XX New peptide keto:amide derivs - are protease inhibitors used to
PT treat e.g. Alzheimer's disease, coagulation disorders and other
PT neurodegenerative disorders.
PT
XX
PS Disclosure; Columns 17-18; 17pp; English.
XX
XX The present sequence represents a peptide ketoamide derivative which is
CC useful for selectively inhibiting cathepsin G. It is a specific example
CC of a new class of peptides which selectively inhibit serine proteases or
CC cysteine proteases, including calpains and cathepsin B. The calpain
CC inhibitors are useful for treatment of various neurodegenerative diseases
CC and conditions including ischaemia, stroke and Alzheimer's disease. The
CC protease inhibitors, especially the elastase, trypsin and chymase
CC inhibitors, are used to control tissue damage and various inflammatory
CC conditions mediated by proteases, such as blistering. They are also
CC useful as anticoagulants and can be used to treat thrombosis. The
CC peptides may also be used to identify new proteolytic enzymes encountered
CC in research. Further, they may also be useful in research and
CC industrially to prevent undesired proteolysis that occurs during the
CC production, isolation, purification, transport and storage of valuable
CC peptides and proteins; e.g. they may be added to antibodies, enzymes,
CC plasma proteins, tissue extracts or other proteins and peptides which are
CC widely sold for use in clinical analyses, biomedical research and for
CC many other reasons.
XX
SQ Sequence 4 AA;

```

Query Match          100.0%; Score 20; DB 18; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
      1111
Db      1 AAPM 4

RESULT 8
AAW26078
ID      AAW26078 standard; peptide; 4 AA.
XX
AC      AAW26078;
XX
DT      05-NOV-1997 (first entry)
XX
DE      Substrate #1 for serine protease.
XX
KW      Serine protease; substrate; Streptomyces griseus; guanidine; pre-soak;
KW      cleaning composition; laundry detergent; additive composition; enzyme;
KW      dishwasher detergent; drain opener; urea; contact lens cleanser;
KW      proteinaceous stain.
XX
OS      Synthetic.
XX
FH      Key Location/Qualifiers
FT      Modified-site 1
FT      Modified-site /note= "succinylated"
FT      Modified-site 4
FT      Modified-site /note= "para-nitroanilidated"
XX
XX      US5646028-A.
PN
PD      08-JUL-1997.
XX
PF      18-JUN-1991; 91US-0718303.
XX
PR      18-JUN-1991; 91US-0718303.
PR      06-NOV-1992; 92US-0973343.
PR      18-AUG-1994; 94US-0292924.
PR      17-OCT-1995; 95US-0544143.
XX
PA      (CLRXX ) CLOROX CO.
PI
PI      Leigh SD;
XX
DR      WPI; 1997-362936/33.
XX
PT      Serine protease from Streptomyces griseus ATCC 55178 - with good
PT      stability in presence of urea or guanidine, useful in cleaning
PT      compositions, including laundry and dishwashing detergents
XX
XX
PS      Example 1; Column 12; 16pp; English.
XX
CC      AAW26078-W26096 represent substrates for the serine protease of the
CC      invention. The serine protease recognises these sequences, but is
CC      specific for the sequence shown in AAW24567. The protease has the
CC      N-terminal and C-terminal sequences represented by AAW24565 and AAW24566
CC      respectively. The serine protease was isolated from Streptomyces griseus
CC      variety alkaliphilus No. 33 (ATCC 55178). The protease has an apparent
CC      molecular weight of 19 kD (by reducing sodium dodecylsulphate
CC      polyacrylamide gel electrophoresis), and improved stability against urea
CC      and guanidine. The protease is inhibited by phenylmethylsulphonyl
CC      fluoride. The serine protease is useful in liquid or granular cleaning
CC      compositions, specifically laundry detergents or additive compositions.
CC      It is also useful in automatic dishwasher detergents, pre-soaks, drain
CC      openers, contact lens cleansers etc. The protease has better activity
CC      against proteinaceous stains than known enzymes and unusually high
CC      stability in the presence of chaotropic agents.
XX
SQ      Sequence 4 AA:
```

```

Query Match          100.0%; Score 20; DB 18; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
      1111
Db      1 AAPM 4

RESULT 9
AAW51453
ID      AAW51453 standard; peptide; 4 AA.
XX
AC      AAW51453;
XX
DT      02-SEP-1998 (first entry)
XX
DE      Indicator for detection of leukocyte esterase activity in urine.
XX
KW      Urine; indicator; detection; leukocyte esterase; activity;
KW      white blood cell.
XX
OS      Synthetic.
XX
FH      Key Location/Qualifiers
FT      Modified-site 1
FT      Modified-site /note= "Methoxysuccinyl-suc-Ala"
FT      Modified-site 4
FT      Modified-site /note= "Met-nitroanillide"
XX
XX      US5776780-A.
PN
PD      07-JUL-1998.
XX
PF      12-APR-1996; 96US-0631581.
XX
PR      12-APR-1996; 96US-0631581.
PR      28-MAY-1993; 93US-0068956.
PR      24-APR-1995; 95US-0429292.
XX
PA      (CHIM-) CHIMERA RES & CHEM INC.
PI
PI      Carter JM, Smith JV;
XX
DR      WPI; 1998-398049/34.
XX
PT      Automated measurement of white blood cell esterase activity in urine
PT      - comprises adding reagent composition to the sample containing an
PT      indicator, placing sample in a analyser and comparing absorbance
PT      with a standard measurement
XX
PS      Claim 11; Column 14; 8pp; English.
XX
XX
CC      The invention relates to a method for the measurement of white blood
CC      cell esterase activity in urine. It comprises: (i) placing an aliquot of
CC      the urine in an automated analyser sampling cup; (ii) placing the cup in
CC      a sampling tray with the automated analyser, transferring the urine to a
CC      cuvette and injecting at least one reagent composition in an aqueous
CC      medium into the cuvette, where the reagent composition comprises a
CC      buffer to adjust the pH of the urine to 7-9, an activator and surfactant
CC      selected from benzalkonium chloride, 2,5-dimethylbenzenesulphonic acid
CC      and benzethonium chloride, and an indicator to determine leukocyte
CC      esterase activity in the urine; (iii) reading the aliquot of urine at
CC      specified intervals, in accordance with a preprogrammed code introduced
CC      into the automated analyser, at a preprogrammed monochromatically
CC      specified wavelength, to compare absorbance of the urine sample and
CC      reagent composition complex with that of a standard containing a known
CC      concentration of leukocyte esterase and determining the quantitative
CC      amount of leukocyte esterase in the patient's urine. The method provides
CC      a fully automated method for determining white blood cells in urine. The
CC      method is preferable to the old dipstick method as it reduces consumable
CC      material and labour costs. It also offers increased accuracy, sensitivity
```

CC and reduction of interference by substances which affected prior art
CC tests. The present sequence represents a specifically claimed indicator.
XX
SQ Sequence 4 AA;
Query Match 100.0%; Score 20; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPM 4
Db 1 AAPM 4
RESULT 10
AAW51449
ID AAW51449 standard; peptide; 4 AA.
XX
AC AAW51449;
XX
DT 02-SEP-1998 (first entry)
XX
DE Indicator for detection of leukocyte esterase activity in urine.
XX
KW Urine; indicator; detection; leukocyte esterase; activity;
XX white blood cell.
XX
OS Synthetic.
XX
FH Key
FT Modified-site 1 Location/Qualifiers
FT Modified-site 4 /note= "Methoxysuccinyl-suc-Ala"
FT Modified-site 4 /note= "Met-nitroanilide"
XX
XX
PN US5776780-A.
XX
PD 07-JUL-1998.
XX
PF 12-APR-1996; 96US-0631581.
XX
PR 12-APR-1996; 96US-0631581.
PR 28-MAY-1993; 93US-0068956.
PR 24-APR-1995; 95US-0429292.
XX
PA (CHIM-) CHIMERA RES & CHEM INC.
XX
PI Carter JM, Smith JV;
XX
DR WPI; 1998-398049/34.
XX
XX
PT Automated measurement of white blood cell esterase activity in urine
PT - comprises adding reagent composition to the sample containing an
PT indicator, placing sample in a analyser and comparing absorbance
PT with a standard measurement
XX
XX
PS Claim 11; Column 14; 8pp; English.
XX
XX The invention relates to a method for the measurement of white blood
CC cell esterase activity in urine. It comprises: (i) placing an aliquot of
CC the urine in an automated analyser sampling cup; (ii) placing the cup in
CC a sampling tray with the automated analyser, transferring the urine to a
CC cuvette and injecting at least one reagent composition in an aqueous
CC medium into the cuvette, where the reagent composition comprises a
CC buffer to adjust the pH of the urine to 7-9, an activator and surfactant
CC selected from benzalkonium chloride, 2,5-dimethylbenzenesulphonic acid
CC and benzethonium chloride, and an indicator to determine leukocyte
CC esterase activity in the urine; (iii) reading the aliquot of urine at
CC specified intervals, in accordance with a preprogrammed code introduced
CC into the automated analyser, at a preprogrammed monochromatically
CC specified wavelength, to compare absorbance of the urine sample and
CC reagent composition complex with that of a standard containing a known
CC concentration of leukocyte esterase and determining the quantitative

CC amount of leukocyte esterase in the patient's urine. The method provides
CC a fully automated method for determining white blood cells in urine. The
CC method is preferable to the old dipstick method as it reduces consumable
CC material and labour costs. It also offers increased accuracy, sensitivity
CC and reduction of interference by substances which affected prior art
CC tests. The present sequence represents a specifically claimed indicator.
XX
SQ Sequence 4 AA;
Query Match 100.0%; Score 20; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPM 4
Db 1 AAPM 4
RESULT 11
AAW51611
ID AAW51611 standard; peptide; 4 AA.
XX
AC AAW51611;
XX
DT 03-SEP-1998 (first entry)
XX
DE Peptide conjugated to lipid for use in liposomal drug delivery.
XX
KW Liposome; conjugate; drug delivery; peptidase-secreting cell;
KW tumour; diagnosis; therapy.
XX
OS Synthetic.
XX
PN WO9816240-A1.
XX
PD 23-APR-1998.
XX
PF 15-OCT-1997; 97WO-US18538.
XX
PR 15-OCT-1996; 96US-0027544.
XX
PA (LIPO) LIPOSOME CO INC.
XX
PI Ali S, Cabralilly D, Erukulla RK, Franklin JC;
PI Janoff AS, Meers PR, Pak C;
XX
XX
DR WPI; 1998-261025/23.
XX
XX
PT New peptide-lipid conjugates are incorporated into liposome(s) - to
PT selectively destabilise the liposome(s) in the vicinity of target
PT peptidase-secreting cells, e.g., tumour cells; useful in diagnosis
PT and therapy
XX
XX
PS Claim 6; Page 33; 55pp; English.
XX
XX The invention relates to peptide-lipid conjugates and their use in
CC the preparation of liposomes which are predisposed to degradation in the
CC presence of peptidase-secreting cells and hence are targeted to these
CC cells. The liposomes can be used to treat mammalian diseases, disorders
CC and conditions, e.g. tumours, microbial infections and inflammation. The
CC liposomes allow selective delivery of an active agent to desired cells.
CC The liposomes are stable when the peptide remains conjugated to the
CC lipid. However, once the peptide portion of the conjugate is cleaved
CC from the lipid, by the action of cell-secreted peptidases, the liposomes
CC destabilise and release their contents in the vicinity of, or into, the
CC characterised by the occurrence of peptidase-secreting cells. For
CC example, the breast cancer cell line MCF-7 is known to secrete elastase,
CC the levels of which are inversely correlated to overall survival in
CC breast cancer patients. The present sequence represents a specifically
CC claimed peptide which forms part of the peptide-lipid conjugate.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
1111
DB 1 AAPM 4

RESULT 12
AAY31239
ID AAY31239 standard; peptide; 4 AA.

AC AAY31239;
XX
DT 09-NOV-1999 (first entry)

DE Alpha-aminoalkyl phosphonate derivative peptide 1.

XX Alpha-aminoalkyl phosphonate; serine protease inhibitor; anticoagulant;
KW cytostatic; anti-inflammatory; gastrointestinal; endocrine; respiratory;
KW dermatological; chymotrypsin inhibitor; elastase inhibitor; inflammatory;
KW trypsin inhibitor; trypsin inhibitor; blood coagulation; pancreatitis;
KW tumour invasion control; emphysema; respiratory distress syndrome;
KW skin blistering.

XX Synthetic.

XX US5952307-A.

XX 14-SEP-1999.

XX 14-AUG-1997; 97US-0907840.

XX 14-AUG-1997; 97US-0907840.

PR 21-JAN-1994; 94US-0184286.

XX (GEOR-) GEORGIA TECH RES CORP.

PI Jackson DS, Ni L, Powers JC;

DR WPI; 1999-539191/45.

XX New aminoalkyl phosphonate derivatives are serine protease
PT inhibitors useful for decreasing blood coagulation and for treating
PT tumour invasion, inflammation, pancreatitis, emphysema, respiratory
PT distress syndrome and skin blistering

PS Disclosure; Column 29-30; 18pp; English.

XX This invention describes the production of novel alpha-aminoalkyl
CC phosphonate derivatives (I) which have anticoagulant, cytostatic,
CC anti-inflammatory, gastrointestinal, endocrine, respiratory and
CC dermatological activity. The products of the invention are also found
CC to act as serine protease inhibitors, chymotrypsin inhibitors, elastase
CC inhibitors, trypsin inhibitors and trypsin inhibitors. (I) are useful
CC for decreasing blood coagulation, and for controlling tumour invasion,
CC inflammations, pancreatitis, emphysema, respiratory distress syndrome
CC and skin blistering. AAY31239-Y31241 are peptides used to describe the
CC method of the invention.

XX Sequence 4 AA;

Query Match 100.0%; Score 20; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
1111
DB 1 AAPM 4

RESULT 13
AAW82444
ID AAW82444 standard; peptide; 4 AA.

XX AAW82444;

XX 24-FEB-1999 (first entry)

DE Seq ID 13 from US5849560.

XX Amyloid beta-protein precursor; endoprotease; human; brain; screening;
KW Alzheimer's disease; O-phenanthroline; metal chelator; treatment;
KW pheyethylsulphonyl fluoride; protease inhibitor.

XX Unknown.

XX US5849560-A.

XX 15-DEC-1998.

XX 26-FEB-1993; 93US-0025321.

XX 26-FEB-1993; 93US-0025321.

PR 17-AUG-1990; 90US-0568806.

PR 05-APR-1991; 91US-0681093.

XX (UYBO-) UNIV BOSTON.

XX Abraham CR;

XX WPI; 1999-069739/06.

XX Purified endoprotease associated with Alzheimer's disease - is
PT prepared from fractions of brain tissue homogenate and is useful for
PT drug screening

XX Disclosure; Column 23-24; 27pp; English.

XX This invention related to the discovery of a human amyloid beta-protein
CC precursor N-terminal domain which is cleaved by a purified endoprotease
CC from human brain tissue homogenate. This endoprotease is identical to
CC an endoprotease found in the brains of humans with Alzheimer's disease.
CC The endoprotease is inhibited by O-phenanthroline and by metal chelators
CC and is not inhibited by pheyethylsulphonyl fluoride. The endoprotease is
CC useful to screen for protease inhibitors that might be useful for
CC treating Alzheimer's disease by inhibiting cleavage of the N-terminal
CC domain of amyloid beta-protein precursor.
CC NOTE: This sequence does not appear in the specification but is
CC represented in the Sed ID listing.

XX Sequence 4 AA;

Query Match 100.0%; Score 20; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
1111
DB 1 AAPM 4

RESULT 14
AAB97634
ID AAB97634 standard; peptide; 4 AA.

XX AAB97634;

XX 21-SEP-2001 (first entry)

DE Alpha-ketoamide cathepsin G inhibitor peptide.

XX Peptide ketoamide; cysteine protease inhibitor;

KW serine protease inhibitor; transition state analogue;

KW neurodegenerative disease; ischaemia; stroke; Alzheimer's disease;
 KW anticoagulant; thrombosis; adult respiratory distress syndrome;
 KW emphysema; rheumatoid arthritis; pancreatitis; viral infection;
 KW muscular dystrophy; myocardial tissue damage; tumour metastasis;
 KW bone resorption; cathepsin inhibitor.

OS Synthetic.

FH Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site 4 /note= "Meo-Suc-Ala"
 FT Modified-site 4 /note= "Met-CO-NR3R4"

XX US6235929-B1.

XX 22-MAY-2001.

PF 27-DEC-1996; 96US-0777354.

XX 27-DEC-1991; 91US-0815073.

PR 09-SEP-1993; 93US-0118997.

PR 20-MAY-1994; 94US-0246511.

PR 06-OCT-1995; 95US-0539944.

XX (GEOR-) GEORGIA TECH RES CORP.

PI Powers JC;

DR WPI; 2001-440210/47.

XX New tripeptide ketoamide derivatives are serine and cysteine protease
 PT inhibitors, useful as anticoagulants and for treating neurodegenerative
 PT diseases, thrombosis, emphysema and rheumatoid arthritis

PS Disclosure; Column 18; 24pp; English.

CC The invention relates to peptide ketoamide derivatives useful for
 CC selectively inhibiting serine proteases and cysteine proteases, and for
 CC generally inhibiting all members of these classes of enzymes. The peptide
 CC ketoamides are derivatives of substrate peptides of serine and cysteine
 CC proteases and act as transition state analogues. The peptides of the
 CC invention have the formula M1-AA1-AA2-AA3-CO-NR3R4 where:

CC M1 is NH2CO-, NH2CS-, NH2SO2-, XNHSO2-, XNHSO-, X2NCS-,
 CC XNHSO2-, X2NSO2-, XCO-, XCS-, XSO2-, XOCO- or XOCS;

CC X is 1-10C (fluoro)alkyl (optionally substituted by J), 1-adamantyl,
 CC 9-fluorenyl, phenyl or naphthyl (optionally substituted by up to 3 of
 CC K), or 1-10C alkylphenyl, 1-10C alkyldiphenyl or 1-10C alkylphenoxy

CC (all optionally substituted by K);

CC J is halo, COOH, OH, CN, NO2, NH2, 1-10C alkoxy, 1-10C alkylamine,
 CC 2-12C dialkylamine, 1-10C alkyl-OCO-, 1-10C alkyl-CONH- or 1-10C
 CC alkylthio;

CC K is halo, 1-10C (perfluoro)alkyl, 1-10C alkoxy, NO2, CN, OH, COOH,
 CC NH2, 1-10C alkylamino, 2-12C dialkylamino, 1-10C acyl, 1-10C alkoxy-CO
 CC or 1-10C alkylthio;

CC AA1, AA2 are independently Ala, Val, Leu, Ile, Gly, Ser, Asp or Glu;
 CC AA3 is Asp or Glu, in either the L or D form;

CC R3 is 2-3C alkylphenyl, 3-20C cycloalkylphenyl, 1-20C alkylphenyl
 CC (substituted by up to 3 of K), 3-20C cycloalkylphenyl (substituted by
 CC K), NH-CH2CH2-(4-hydroxyphenyl) or NH-CH2CH2-(3-indolyl); and

CC R4 = H, 3-20C alkyl, cycloalkyl, 1-20C alkylphenyl (optionally
 CC substituted by up to 3 of K), 3-20C cycloalkylphenyl (optionally
 CC substituted by K), NHCH2CH2-(4-hydroxyphenyl) or

CC NHCH2CH2-2-(3-indolyl).

CC The peptide ketoamides are useful for treating neurodegenerative diseases
 CC (including ischaemia, stroke and Alzheimer's disease), as anticoagulants,
 CC and for treating thrombosis. They are also useful for treating

CC emphysema, adult respiratory distress syndrome, rheumatoid arthritis,
 CC pancreatitis, viral infections, muscular dystrophy, myocardial tissue
 CC damage, tumour metastasis and bone resorption. The present sequence
 CC represents a cathepsin G (cysteine protease) peptide inhibitor.

XX Sequence 4 AA;

Query Match 100.0%; Score 20; DB 22; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AAPM 4
 Db 1 AAPM 4

RESULT 15
 AAB82199

ID AAB82199 standard; Peptide; 4 AA.

XX AC AAB82199;

DT 24-JUL-2001 (first entry)

DE Cathepsin G recognition site #1.

XX Oligomeric ligand display; fusion protein; viral surface;

KW drug screening; factor Xa; elastase recognition site.

XX Synthetic.

FH Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site 4 /label= Meo-Suc-Ala
 FT Modified-site 4 /label= Met-pNA

XX WO200127292-A2.

PD 19-APR-2001.

PF 09-OCT-2000; 2000WO-GB03858.

PR 08-OCT-1999; 99US-0415565.

PA (CDDH-) CDD HOLDINGS LTD.

PI Chadwick MP, Belcher C, Glenn D, Bullough FJ, Russell SJ;

DR WPI; 2001-282039/29.

PT Recombinant fusion protein for screening for compounds that can
 PT modulate ligand/receptor binding and for delivering a transferable
 PT label to a target cell, comprises two polypeptides, one of which forms
 PT a oligomeric ligand

PS Disclosure; Page 13; 45pp; English.

CC The present invention relates to a method for displaying oligomeric
 CC ligands as recombinant fusion proteins on viral surfaces. The fusion
 CC proteins comprise two polypeptides fused to each other through peptide
 CC bonding, where the first polypeptide comprises two domains in a
 CC continuous polypeptide chain which self-associate to form a
 CC homo-oligomeric ligand or an oligomeric ligand comprising a disulfide
 CC bonded growth factor, which has the ability to bind to a receptor. The
 CC method of the present invention is useful for drug screening. Optionally,
 CC a protease recognition site can be present between the first and second
 CC polypeptides. The present sequence is one such protease recognition site
 CC which can be used in the method of the present invention.

XX Sequence 4 AA;

Query Match 100.0%; Score 20; DB 22; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
 Db 1 AAPM 4

| | |
|-----------------------|--|
| RESULT | 16 |
| AAB62788 | |
| ID | AAB62788 standard; peptide; 4 AA. |
| XX | |
| AC | AAB62788; |
| DT | 03-APR-2001 (first entry) |
| XX | |
| DE | Peptide-lipid conjugate peptide #3. |
| XX | |
| KW | Peptide-lipid conjugate; liposome; liposomal drug delivery; cancer; |
| KW | peptidase-secreting cell; phosphatidylethanolamine; |
| XX | inflammatory disorder; neuropathy. |
| OS | Synthetic. |
| XX | |
| FH | Key |
| FT | Modified-site |
| FT | 1 |
| FT | /label= OTHER |
| FT | /note= "optionally modified by succinyl, methoxysuccinyl or carboxy sugar group" |
| FT | 4 |
| FT | Modified-site |
| FT | 4 |
| FT | /label= OTHER |
| FT | /note= "C-terminal phosphatidylethanolamine" |
| XX | |
| XX | WO200100247-A1. |
| PN | |
| PD | 04-JAN-2001. |
| XX | |
| XX | 13-JUN-2000; 2000WO-US16248. |
| PE | |
| XX | 29-JUN-1999; 99US-0343650. |
| PR | |
| XX | |
| PA | (LIPO) LIPOSOME CO INC. |
| XX | |
| PI | Meers P, Pak C, Ali S, Janoff A, Franklin JC, Erukulla R; |
| PI | Cabral-Lilly D, Ahl P; |
| DR | WPI; 2001-122976/13. |
| XX | |
| LIP | Liposomal drug delivery for treating cancer, inflammatory, genetic disorders and microbial infections, involves administering liposomes comprising peptide-lipid conjugates - |
| PT | |
| PS | Claim 4; Page 61; 107pp; English. |
| XX | |
| CC | The present invention describes a method of administering the contents of a liposome to a cell by administering a liposome containing a peptide-lipid conjugate where the peptide is cleavable by a peptidase and the lipid is a phosphatidylethanolamine. This can be used in liposomal drug delivery, for example in the treatment of cancer, inflammatory disorders and neuropathies. |
| CC | |
| CC | |
| Sequence | 4 AA; |
| Query Match | 100.0%; Score 20; DB 22; Length 4; |
| Best Local Similarity | 100.0%; Pred. No. 7.8e+05; |
| Matches | 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| OY | 1 AAPM 4 |
| | |
| DB | 1 AAPM 4 |
| RESULT | 17 |
| AAB83504 | |
| ID | AAB83504 standard; peptide; 4 AA. |
| XX | |
| AC | AAB83504; |
| XX | |
| DT | 18-SEP-2002 (first entry) |

| | | |
|----|---|--|
| XX | | Enzyme cleavable peptide to project from lipid-based carrier #4. |
| DE | | |
| XX | | Cytostatic; antibacterial; anti-inflammatory; antifungal; |
| KW | tumour; cancer; microbial infection; inflammatory disorder; | |
| KW | liposome. | |
| XX | | |
| OS | Synthetic. | |
| PN | US6339069-B1. | |
| XX | | |
| PD | 15-JAN-2002. | |
| XX | | |
| PF | 29-JUN-1999; | 99US-0343650. |
| XX | | |
| PR | 15-OCT-1996; | 96US-027544P. |
| PR | 27-FEB-1997; | 97US-039183P. |
| PR | 15-OCT-1997; | 97US-0950618. |
| PR | 07-OCT-1998; | 98US-0168010. |
| XX | | |
| PA | (ELAN-) ELAN PHARM TECHNOLOGIES INC. | |
| XX | | |
| PI | Meers PR, Pak C, Ali S, Janoff A, Franklin JC, Erukulla RK; | |
| PI | Cabral-Lilly D, Ahl PL; | |
| XX | | |
| DR | WPI; 2002-266419/31. | |
| XX | | |
| PT | Liposome composition useful in the treatment of e.g. tumours comprises | |
| PT | liposome containing a bioactive agent and a lipid component which is | |
| PT | delivered to the vicinity of cells in the mammal | - |
| PS | Claim 4; Column 8; 50pp; English. | |
| XX | | |
| CC | This invention relates to a method of administering a bioactive agent | |
| CC | to a mammal by administering a composition which comprises a carrier | |
| CC | and a liposome. The liposome is delivered to the vicinity of cells in | |
| CC | the mammal which secrete a peptidase which recognizes the amino acid | |
| CC | residue. The composition in considered cytostatic, antibacterial | |
| CC | anti-inflammatory and antifungal. The composition is used to treat | |
| CC | animals for diseases such as tumours, cancer, microbial infection or | |
| CC | inflammatory disorders. The peptide-lipid conjugate selectively | |
| CC | destabilising the liposomes in the vicinity of target peptidase-secreting | |
| CC | cells and hence delivers the liposomes to the vicinity of the target | |
| CC | cells or directly into the cells. The present sequence represents | |
| CC | an enzyme cleavable peptide which is of sufficient length to project | |
| CC | above the lipid-based carriers of which they have been incorporated. | |
| XX | | |
| SO | Sequence | 4 AA; |
| | Query Match | 100.0%; Score 20; DB 23; Length 4; |
| | Best Local Similarity | 100.0%; Pred. No. 7.8e+05; |
| | Matches | 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0. |
| OY | 1 AAPM 4 | |
| | | |
| DB | 1 AAPM 4 | |
| | | |
| | RESULT .18 | |
| | ABB83518 | |
| ID | ABB83518 standard; peptide; 4 AA. | |
| XX | | |
| AC | ABB83518; | |
| XX | | |
| DT | 18-SEP-2002 (first entry) | |
| XX | | |
| DE | Enzyme cleavable peptide to project from lipid-based carrier #19. | |
| XX | | |
| KW | Cytostatic; antibacterial; anti-inflammatory; antifungal; | |
| KW | tumour; cancer; microbial infection; inflammatory disorder; | |
| KW | liposome. | |
| XX | | |
| OS | Synthetic. | |

XX Key Location/Qualifiers
FH Modified-site 1
FT /label= MeoSuc
FT /note= "Methoxy succinyl"
XX
XX US6339069-B1.
XX
XX 15-JAN-2002.
XX
XX 29-JUN-1999; 99US-0343650.
XX
XX 15-OCT-1996; 96US-027544P.
XX 27-FEB-1997; 97US-039183P.
XX 15-OCT-1997; 97US-0950618.
XX 07-OCT-1998; 98US-0168010.
XX
XX (ELAN-) ELAN PHARM TECHNOLOGIES INC.
XX
XX Meers PR, Pak C, Ali S, Janoff A, Franklin JC, Erukulla RK;
XX Cabral-Lilly D, Ahl PL;
XX
XX WPI; 2002-266419/31.
XX
XX Liposome composition useful in the treatment of e.g. tumours comprises
XX liposome containing a bioactive agent and a lipid component which is
XX delivered to the vicinity of cells in the mammal -
XX
XX Disclosure; Column 15; 50pp; English.
XX
XX This invention relates to a method of administering a bioactive agent
XX to a mammal by administering a composition which comprises a carrier
XX and a liposome. The liposome is delivered to the vicinity of cells in
XX the mammal which secrete a peptidase which recognizes the amino acid
XX residue. The composition in considered cytostatic, antibacterial
XX anti-inflammatory and antifungal. The composition is used to treat
XX animals for diseases such as tumours, cancer, microbial infection or
XX inflammatory disorders. The peptide-lipid conjugate selectively
XX destabilising the liposomes in the vicinity of target peptidase-secreting
XX cells and hence delivers the liposomes to the vicinity of the target
XX cells or directly into the cells. The present sequence represents
XX an enzyme cleavable peptide which is of sufficient length to project
XX above the lipid-based carriers of which they have been incorporated.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 23; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
 ||||
Db 1 AAPM 4

RESULT 19
AAO18044
ID AAO18044 standard; peptide; 4 AA.
XX
XX AAO18044;
AC
XX
DT 30-AUG-2002 (first entry)
XX
XX C-terminal truncated apoE formation inhibitor peptide #2.
DE
XX
KW C-terminal truncated apoE; apolipoprotein E; apoE; Alzheimer's disease;
KW neurofibrillary tangle; apoE4 allele; nootropic; neuroprotective;
KW cardiant; vulnerary; cerebroprotective; coronary artery disease;
KW head trauma; stroke.
XX
OS Synthetic.
XX
XX WO200238108-A2.
PN

XX
PD 16-MAY-2002.
XX
XX 02-NOV-2001; 2001WO-US51172.
PF
XX
XX 03-NOV-2000; 2000US-245737P.
XX
XX (GLAD-) GLADSTONE INST J DAVID.
XX
XX Huang Y, Mahley RW;
PI
XX
XX WPI; 2002-490051/52.
XX
XX
XX Inhibiting neurofibrillary tangles formation, useful for treating e.g.
XX Alzheimer's, coronary artery disease or stroke, by reducing the
XX formation of carboxyl-terminal truncated form of apolipoprotein E in a
XX neuron of the individual -
XX
XX Claim 30; Page 63; 75pp; English.
XX
XX The present invention relates to a method of inhibiting the formation of
XX neurofibrillary tangles in an individual, which involves reducing the
XX formation of a carboxyl-terminal truncated form of apolipoprotein E
XX (apoE) in a neurone in the individual. The method is useful for
XX inhibiting the formation of neurofibrillary tangles in an individual. The
XX reduction in the formation of carboxyl-terminal truncated apoE treats a
XX disorder related to apoE in an individual, specifically Alzheimer's
XX disease, coronary artery disease, head trauma or stroke. The present
XX sequence is a peptide capable of inhibiting the formation of the
XX C-terminal truncated form of apoE.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 23; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
 ||||
Db 1 AAPM 4

RESULT 20
ABB83531
ID ABB83531 standard; peptide; 5 AA.
XX
XX ABB83531;
AC
XX
DT 18-SEP-2002 (first entry)
XX
XX Enzyme cleavable peptide to project from lipid-based carrier #32.
DE
XX
KW Cytostatic; antibacterial; anti-inflammatory; antifungal;
KW tumour; cancer; microbial infection; inflammatory disorder;
KW liposome.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /label= Suc
FT /note= "Succinyl"
XX
XX US6339069-B1.
PN
XX
XX 15-JAN-2002.
PD
XX
PF 29-JUN-1999; 99US-0343650.
XX
XX 15-OCT-1996; 96US-027544P.
XX 27-FEB-1997; 97US-039183P.
XX 15-OCT-1997; 97US-0950618.
XX 07-OCT-1998; 98US-0168010.
PN

XX (ELAN-) ELAN PHARM TECHNOLOGIES INC.
PA Meers PR, Pak C, Ali S, Janoff A, Franklin JC, Erukulla RK;
XX Cabral-Lilly D, Ahl PL;
PI WPI; 2002-266419/31.
XX
DR liposome composition useful in the treatment of e.g. tumours comprises
PT liposome containing a bioactive agent and a lipid component which is
PT delivered to the vicinity of cells in the mammal -
XX
PS Disclosure; Column 15; 50pp; English.
XX This invention relates to a method of administering a bioactive agent
CC to a mammal by administering a composition which comprises a carrier
CC and a liposome. The liposome is delivered to the vicinity of cells in
CC the mammal which secrete a peptidase which recognizes the amino acid
CC residue. The composition in considered cytostatic, antibacterial
CC anti-inflammatory and antifungal. The composition is used to treat
CC animals for diseases such as tumours, cancer, microbial infection or
CC inflammatory disorders. The peptide-lipid conjugate selectively
CC destabilising the liposomes in the vicinity of target peptidase-secreting
CC cells and hence delivers the liposomes to the vicinity of the target
CC cells or directly into the cells. The present sequence represents
CC an enzyme cleavable peptide which is of sufficient length to project
CC above the lipid-based carriers of which they have been incorporated.
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 20; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 2 AAPM 5

RESULT 21
ABB83532
ID ABB83532 standard; peptide; 5 AA.
XX
AC ABB83532;
XX
DT 18-SEP-2002 (first entry)
XX
DE Enzyme cleavable peptide to project from lipid-based carrier #33.
XX
KW Cytostatic; antibacterial; anti-inflammatory; antifungal;
KW tumour; cancer; microbial infection; inflammatory disorder;
KW liposome.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Carboxy sugar"
XX
XX
PN US6339069-B1.
PD 15-JAN-2002.
XX
PF 29-JUN-1999; 99US-0343650.
XX
PR 15-OCT-1996; 96US-027544P.
PR 27-FEB-1997; 97US-039183P.
PR 15-OCT-1997; 97US-0950618.
PR 07-OCT-1998; 98US-0168010.
XX
PA (ELAN-) ELAN PHARM TECHNOLOGIES INC.
XX Meers PR, Pak C, Ali S, Janoff A, Franklin JC, Erukulla RK;
PI

PI Cabral-Lilly D, Ahl PL;
XX
XX WPI; 2002-266419/31.
DR
XX liposome composition useful in the treatment of e.g. tumours comprises
PT liposome containing a bioactive agent and a lipid component which is
PT delivered to the vicinity of cells in the mammal -
XX
PS Disclosure; Column 15; 50pp; English.
XX
CC This invention relates to a method of administering a bioactive agent
CC to a mammal by administering a composition which comprises a carrier
CC and a liposome. The liposome is delivered to the vicinity of cells in
CC the mammal which secrete a peptidase which recognizes the amino acid
CC residue. The composition in considered cytostatic, antibacterial
CC anti-inflammatory and antifungal. The composition is used to treat
CC animals for diseases such as tumours, cancer, microbial infection or
CC inflammatory disorders. The peptide-lipid conjugate selectively
CC destabilising the liposomes in the vicinity of target peptidase-secreting
CC cells and hence delivers the liposomes to the vicinity of the target
CC cells or directly into the cells. The present sequence represents
CC an enzyme cleavable peptide which is of sufficient length to project
CC above the lipid-based carriers of which they have been incorporated.
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 20; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 2 AAPM 5

RESULT 22
AAB06270
ID AAB06270 standard; peptide; 9 AA.
XX
AC AAB06270;
XX
DT 03-OCT-2000 (first entry)
XX
DE Human proBDNF signal peptide cleavage site.
XX
KW Human; pro-brain-derived neurotrophic factor; proBDNF; SKI-1;
KW subtilisin-kexin isoenzyme; antilipemic; cytostatic; vasotropic;
KW SKI-1 inhibitor; hypercholesterolaemia; liver steatosis;
KW Ras-dependent cancer; restenosis; amyloid protein formation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Cleavage-site 5..6
FT
PN WO200026348-A2.
PD 11-MAY-2000.
XX
PF 04-NOV-1999; 99WO-CA01058.
XX
PR 04-NOV-1998; 98CA-2249648.
XX
PA (RECL-) INST RECH CLINIQUES MONTREAL.
XX
PI Seidah N, Chretien M, Marcinkiewicz M, Laaksonen R, Davignon J;
XX WPI; 2000-365601/31.
XX
PT Novel soluble proteic fragment of subtilisin-kexin isoenzyme for
PT producing a polypeptide useful for treating hypercholesterolemia, liver
PT steatosis and amyloidosis, comprises a specific amino acid sequence -
XX

PS Example 1; Page 22; 119pp; English.
XX
CC The present sequence is the site at which human pro-brain-derived
CC neurotrophic factor (PROBDNF) is cleaved to remove the 18 amino acid
CC signal peptide. PROBDNF is a substrate of subtilisin-kexin isoenzyme 1
CC (SKI-1), a type-1 membrane-bound proteinase. SKI-1 cleaves at a specific
CC threonine residue within the N-terminal segment of human PROBDNF.
CC peptides which bind to and are cleaved by SKI-1 may be used for
CC monitoring SKI-1 activity, for screening inhibitors of SKI-1 activity, or
CC for screening enhancers of SKI-1 activity. Proteic fragments of SKI-1
CC which bind to the SKI-1 catalytic site may be used as inhibitors of SKI-1
CC activity. They may be used to treat diseases involving overexpression of
CC SKI-1 or SKI-1 substrate. Such diseases include hypercholesterolaemia,
CC high levels of fatty acids, lipids or farnesyl pyrophosphate, liver
CC steatosis, Ras-dependent cancer, restenosis and amyloid protein
CC formation.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 20; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
|||
Db 5 AAPM 8

RESULT 23
AAB10855
ID AAB10855 standard; Protein; 14 AA.
XX
AC AAB10855;
XX
DT 02-FEB-2001 (first entry)
XX
DE pSectag-VEGF protein fragment VEGF back1.
XX
DE MVP; multivalent protein; treatment; dressing; skin; mucus; musculature;
KW nervous system; inner organ; hematopoietic system; immune system; joint;
KW support tissue; immunization.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO200053790-A1.
XX
PD 14-SEP-2000.
XX
PE 26-FEB-2000; 2000WO-EP01612.
XX
PR 10-MAR-1999; 99DE-1010419.
XX
PA (AVET) AVENTIS PHARMA DEUT GMBH.
XX
PI Kontermann R, Nettelbeck D, Sedlacek H, Mueller R;
XX
DR WPI; 2000-572272/53.
DR N-PSDB; AAA98162.
XX
PT Cell specific multivalent proteins useful for targeting specific cells
PT for the treatment of disease -
XX
PS Example 1; Page 54; 81pp; German.
XX
CC This invention describes a method for the production of a novel cell
CC specific multivalent protein (MVP). The invention also describes (1) a
CC nucleic acid construct, which encodes an MVP; (2) a bacterium, yeast or
CC mammalian cell, in which the nucleic acid construct of (1) is introduced;
CC (3) the MVP bound to a vector; (4) production of an MVP; (5) an MVP
CC comprising a scfv with a binding site for the adenoviral fibre protein or
CC CD3 molecule and two VEBF units, bound by a peptide linker; and (6) a
CC complex comprising at least two MVPs as above (in which each single

CC ligand can be 0=1). The MVP, optionally bound to a vector, is useful for
CC production of a remedy to treat cells outside tissue by dressings for
CC skin, mucus, nervous systems, inner organs, hematopoietic systems,
CC immune systems, musculature, support tissues or joints and to immunize to
CC prevent or treat diseases.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 20; DB 21; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
|||
Db 4 AAPM 7

RESULT 24
AAB98208
ID AAB98208 standard; Peptide; 15 AA.
XX
AC AAB98208;
XX
DT 17-AUG-2001 (first entry)
XX
DE Human P24 protein-22 N-terminal peptide SEQ ID NO:7.
XX
DE Human; P24 protein-22; diagnosis; treatment; cancer; haemopathy; HIV;
KW human immunodeficiency virus; infection; immunological disease;
KW inflammation.
XX
OS Homo sapiens.
XX
PN WO200129076-A1.
XX
PD 26-APR-2001.
XX
PE 16-OCT-2000; 2000WO-CN00332.
XX
PR 18-OCT-1999; 99CN-0116990.
XX
PA (SHAN-) SHANGHAI BIO ROAD GENE DEV LTD.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2001-290896/30.
XX
PT Human P24 protein-22 and encoded polynucleotide, applicable in
PT diagnosis and treatment of cancer, hemopathy, immunological diseases
PT and various inflammations -
XX
PS Example 6; Page.15; 29pp; Chinese.
XX
CC The present invention describes the human P24 protein-22. The P24
CC protein-22 polypeptide and polynucleotide can be used in the diagnosis
CC and treatment of cancer, haemopathy, HIV (human immunodeficiency virus)
CC infection, immunological diseases and various inflammation. The present
CC sequence represents a the human P24 protein-22 N-terminal peptide, which
CC is used in an example from the present invention.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 20; DB 22; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
|||
Db 2 AAPM 5

RESULT 25
AAR42608

ID AAR42608 standard; Protein; 17 AA.
XX AAR42608;
AC
XX
XX 28-OCT-1993 (first entry)
DT
XX
DE Encoded by human VEGF-165 exon II.
XX
XX Angiogenesis; wound healing; mitogen; vascular endothelial cells;
KW Vascular Endothelial Cell Growth Factor; hVEGF-165; hVEGF-121;
KW alternative RNA splicing.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Cleavage-site 4..5 /note= "signal peptide-mature protein junction"
XX
XX US5219739-A.
PN
XX 15-JUN-1993.
PD
XX
PF 27-JUL-1989; 89US-0387545.
XX
XX 27-JUL-1989; 89US-0387545.
PR 14-DEC-1989; 89US-0450883.
PR 27-JUL-1990; 90US-0559041.
XX
XX (SCIO-) SCIOS NOVA INC.
PA
XX
XX Abraham JA, Fiddes JC, Mitchell RL, Tischer EG;
PI
XX WPI: 1993-205302/25.
DR N-PSDB; AAQ49604.
DR
XX
XX Isolated DNA sequences, expression vectors and transformant cells
PT - used for large scale prodn. of vascular endothelial cell growth
PT factor, for treating wounds in which neo-vascularisation is
PT required
XX
XX Claim 8; Fig 8; 40pp; English.
PS
XX The sequences of the 8 possible exons encoding human vascular
CC endothelial cell growth factor, together with contiguous splice
CC junctions, were obtained from overlapping genomic inserts. A method
CC for producing VEGF is claimed comprising culturing mammalian cells
CC transformed with an expression vector containing exons I-V and
CC VIII. See AAQ44261 for exon I and AAQ49604-Q49610 for exons II-VIII.
CC
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 20; DB 14; Length 17;
Best Local Similarity 100.0%; Pred. NO. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 4 AAPM 7

RESULT 26
AAR94034
ID AAR94034 standard; Protein; 17 AA.
XX
XX AAR94034;
AC
XX 10-OCT-1996 (first entry)
DT
XX
DE VEGF exon II.
XX
XX Vascular endothelial growth factor; VEGF; human; conjugate; tumour; iris;
KW proliferation inhibition; VEGF-mediated pathophysiological condition;
KW dermatological disorder; VEGF receptor; vascular proliferation; retina;

KW ophthalmic disorder; hyperproliferating blood vessel; therapy; psoriasis;
KW conjunctiva; vitreous humour; rheumatoid arthritis; skin cancer;
KW varicose veins; gene therapy.
XX
XX Homo sapiens.
OS
XX
XX WO9606641-A1.
PN
XX 07-MAR-1996.
PD
XX
XX 29-AUG-1995; 95WO-US10973.
PF
XX
XX 16-MAY-1995; 95US-0441979.
PR 29-AUG-1994; 94US-0297961.
PR
XX
XX (PRIZ-) PRIZM PHARM INC.
PA
XX
XX Fleurbaaij GA, Freund E, Houston LL, Nova MP, Sosnowski BA;
PI Victor KD;
PI
XX
XX WPI: 1996-160151/16.
DR N-PSDB; AAT17742.
DR
XX
XX Vascular endothelial cell growth factor (VEGF) conjugates - having
PT VEGF linked to targeted agent, used for inhibiting proliferation of
PT cells, e.g. for gene therapy
PT
XX
XX Disclosure; Page 118; 193pp; English.
PS
XX
XX AAR94033-R94038, AAR94041, AAR94042 and AAW00582 represent vascular
CC endothelial growth factors (VEGF) exons. This sequence represents exon
CC II. These sequences were used in VEGF conjugates of the invention. In
CC the conjugates, VEGF (or fragments of it) are linked to a targeted agent
CC (this can be via a linker sequence), so that the conjugate binds to a
CC VEGF receptor. Cys-modified forms of VEGF are particularly suitable for
CC chemical conjugation to linkers and targeted agents. The conjugates are
CC used for inhibiting proliferation of cells bearing VEGF receptors. They
CC can be used for treating a VEGF-mediated pathophysiological condition,
CC including dermatological disorders with underlying vascular
CC proliferation, solid tumours or an ophthalmic disorder of
CC hyperproliferating blood vessels of the retina, iris, conjunctiva or
CC vitreous humour. The conjugates can also be used for treating
CC psoriasis, rheumatoid arthritis, skin cancers and other tumours, or
CC varicose veins. They are also suitable for use in gene therapy.
CC
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 20; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. NO. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 4 AAPM 7

RESULT 27
AAU00655
ID AAU00655 standard; Peptide; 19 AA.
XX
XX AAU00655;
AC
XX 07-SEP-2001 (first entry)
DT
XX
XX Human targeting peptide sequence #1.
DE
XX
XX Membrane translocating peptide; MTRP; human; intracellular gene delivery;
KW epithelial cell layer; gastrointestinal tract; circulatory system;
KW targeting peptide.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH

```
FT Modified-site 1
FT /note= "N-terminus has a dansyl group"
FT Modified-site 19
FT /note= "C-terminal amide"
XX
XX WO200127154-A2.
XX
XX 19-APR-2001.
XX
XX 27-SEP-2000; 2000WO-IB01491.
XX
XX 27-SEP-1999; 99US-0156246.
XX
XX (OMAH/) O'MAHONY D J.
XX (LAMB/) LAMBKIN I J.
XX
XX O'mahony DJ, lambkin IJ;
XX
XX WPI; 2001-300212/31.
XX
XX Compositions for enhancing uptake of e.g. drugs or DNA across a cell
XX membrane, comprise membrane translocating peptides having specific
XX amino acid sequences or a derivative, fragment, motif, analog or
XX peptidomimetic of the peptides -
XX
XX Example 1; Page 22; 42pp; English.
XX
XX The sequence represents a human targeting peptide which is used to
XX identify the presence of a human membrane translocated peptide (MTLP) and
XX quantify the amount, to bind the MTLP to the surface of a particle, or to
XX localise the MTLP in a cell or tissue sample. MTLPs and their related
XX fragments, motifs, derivatives and analogues are used for enhancing
XX uptake of a pharmaceutically active agent into a cell, into or out of an
XX intracellular compartment and across a cell layer (for example, an
XX epithelial cell layer lining the gastrointestinal tract), either directly
XX or from a pharmaceutically active agent loaded particle, into the
XX circulatory system of an animal. This method is useful for intracellular
XX gene delivery, as a rapid screening method for the identification of
XX MTLPs which retain the functional activity of a full-length MTLP, as a
XX cell-based screen for assaying the functional activity of a MTLP and
XX characterising the properties of a MTLP, for diagnosis of a pathological
XX disorder (by administration of a MTLP-active agent complex or MTLP-active
XX particle complex comprising a diagnostic agent) and for preventing or
XX treating a pathological disorder.
XX
XX Sequence 19 AA;
SQ
XX
XX Query Match 100.0%; Score 20; DB 22; Length 19;
XX Best Local Similarity 100.0%; Pred. No. 2.6e+02;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AAPM 4
XX ||||
DB 13 AAPM 16
XX
XX
XX RESULT 28
XX ABG62530
XX ID ABG62530 standard; Peptide; 25 AA.
XX
XX AC ABG62530;
XX
XX 21-AUG-2002 (first entry)
XX
XX Eubacterial MutS1 DNA polymerase III beta subunit binding peptide #35.
XX
XX DNA polymerase III; beta subunit; eubacteria; antibacterial;
XX eubacterial infection.
XX
XX Burkholderia cepacia.
XX
XX WO200238596-A1.
XX
```

```
PD 16-MAY-2002.
XX
XX 08-NOV-2001; 2001WO-AU01436.
XX
XX 08-NOV-2000; 2000AU-0001320.
XX 06-FEB-2001; 2001AU-0002919.
XX
XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX
XX Dalrymple BP, Kongsuwan K, Wiffels GL, Jennings PA, Kemp GW;
XX
XX WPI; 2002-471546/50.
XX
XX
XX New molecule having surface analogous to surface of domain of
XX eubacterial beta protein contacted by proteins that interact with beta
XX protein, useful to identify inhibitors of beta protein-ligand
XX interaction -
XX
XX Example 1; Page 31; 326pp; English.
XX
XX The invention relates to a molecule (I) comprising a surface (S)
XX analogous to the surface of the domain of eubacterial beta protein
XX contacted by proteins that interact with beta protein, where the
XX surface is defined by the residues X(170), X(172), X(175), X(177),
XX X(241), X(242), X(247), X(346), X(360), and X(362), where the
XX superscript numbers designate the position of residues in Escherichia
XX coli beta protein, or the equivalent residues in homologues from other
XX species of eubacteria, and where:
XX X(170) = Val, Ile, Ala, Thr, Ser or Glu; X(172) = Thr, Ser or Ile;
XX X(175) = His, Tyr, Phe, Lys, Ile, Gln or Arg; X(177) = Leu, Met, Ile,
XX Phe, Val or Ala; X(241) = Phe, Tyr or Leu; X(242) = Pro, Leu or Ile;
XX X(247) = Val, Ile, Ala, Phe, Leu or Met; X(346) = Ser, Pro, Ala, Tyr or
XX Lys; X(360) = Ile, Leu or Val; and X(362) = Met, Leu, Val, Ser, Thr or
XX Arg. Also included are methods of identifying a modulator of the
XX interaction between a eubacterial beta protein and proteins that interact
XX with them, reducing (M4) the effect of eubacterial infestation of a
XX biological system, involves delivering to a system infested with a
XX eubacterial species, a modulator of the interaction between eubacterial
XX beta protein and proteins that interact with the beta protein; and
XX (4) a template (II) for the design of a compound that binds to at least
XX part of (S) of beta protein as defined above comprises a (P) such as
XX X1X2, X3X1X2, X3X1X2X4, Glx1X5X3X1X2, Glx1X5X6X3X6,
XX where: x = any amino acid residue; X1 = Leu, Met, Ile, or Phe;
XX X2 = Leu, Ile, Val, Cys, Phe, Tyr, Trp, Pro, Asp, Ala or Gly;
XX X3 = Ala, Gly, Thr, Asn, Asp, Ser, or Pro; X4 = Ala or Gly; X5 = Leu;
XX and X6 = Leu, Ile, Val, Cys, Phe, Tyr, Trp or Pro. The method are useful
XX for identifying a modulator of the interaction between a eubacterial beta
XX protein and proteins that interact with the beta protein. (M4) is useful
XX for reducing the effect of eubacterial infestation of a biological
XX system. The compounds identified using above mentioned methods are
XX useful as antibacterial agent for treatment or prevention of disease in
XX humans, animals and plants. The present sequence is a eubacterial
XX peptide from a DNA binding protein or polymerase which contains a DNA
XX polymerase III beta subunit binding site.
XX
XX Sequence 25 AA;
SQ
XX
XX Query Match 100.0%; Score 20; DB 23; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 3.3e+02;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AAPM 4
XX ||||
DB 16 AAPM 19
XX
XX
XX RESULT 29
XX AAM19115
XX ID AAM19115 standard; Protein; 31 AA.
XX
XX AC AAM19115;
XX
XX 12-OCT-2001 (first entry)
XX
```

XX Peptide #5549 encoded by probe for measuring cervical gene expression.
DE
XX
KW Probe: human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 27; SEQ ID No 23941; 487bp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP: see AAI10068-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 31 AA;

Query Match 100.0%; Score 20; DB 22; Length 31;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
 ||||
Db 14 AAPM 17

RESULT 30
AAM31764
ID AAM31764 standard; Protein: 31 AA.
XX
AC AAM31764;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #5801 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX

PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 27; SEQ ID No 32033; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 31 AA;

Query Match 100.0%; Score 20; DB 22; Length 31;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
 ||||
Db 14 AAPM 17

RESULT 31
ABG41275
ID ABG41275 standard; Peptide: 31 AA.
XX
AC ABG41275;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 30940.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX
PS Claim 27; SEQ ID No 30940; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 31 AA;
QY
Query Match 100.0%; Score 20; DB 23; Length 31;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPM 4
Db 14 AAPM 17
RESULT 32
AAB39021
ID AAB39021 standard; Protein; 42 AA.
XX
AC AAB39021;
XX

DT 02-FEB-2001 (first entry)
XX
DE Human secreted protein BLAST search protein SEQ ID NO: 111.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200056880-A1.
XX
PD 28-SEP-2000.
XX
PF 16-MAR-2000; 2000WO-US06781.
XX
PR 19-MAR-1999; 99US-0125363.
PR 08-DEC-1999; 99US-0169617.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-602220/57.
XX
PT Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating disorders such as Parkinson's and
PT Alzheimer's diseases, cancers and infections -
XX
PS Disclosure; Page 390; 422pp; English.
XX
CC The invention relates to the isolation of genes AAC59679-C59728 encoding
CC the human secreted proteins AAB38971-B39020. This sequence represents a
CC peptide fragment homologous to the protein encoded by the gene isolated
CC in the present invention. The sequence is a search result from a BLASTX
CC homology search. The genes and proteins are useful for preventing,
CC ameliorating or treating medical conditions, e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such
CC as myocardial ischaemias; (d) wound healing; (e) neurological diseases
CC e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such
CC as viral, bacterial, fungal and parasitic infections.
XX
SQ Sequence 42 AA;
QY
Query Match 100.0%; Score 20; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPM 4
Db 18 AAPM 21
RESULT 33
ABG27642
ID ABG27642 standard; Protein; 43 AA.
XX
AC ABG27642;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #27633.
XX


```
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS91829.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX Claim 20; SEQ ID No 58001; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 43 AA;
SQ
Query Match 100.0%; Score 20; DB 22; Length 43;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AAPM 4
DB 25 AAPM 28
RESULT 34
AAU64697
ID AAU64697 standard; Protein; 50 AA.
XX
XX AAU64697;
AC
XX
XX 27-FEB-2002 (first entry)
DT
XX
XX Propionibacterium acnes immunogenic protein #25593.
DE
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
```

```
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US12865.
XX
XX 21-APR-2000; 2000US-199047P.
XX 02-JUN-2000; 2000US-208841P.
XX 07-JUL-2000; 2000US-216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX N-PSDB; AAS59648.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX vaccinating against and diagnosing infections, especially useful for
XX treating acne vulgaris -
XX
XX Example 1; SEQ ID No 25892; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA).
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 50 AA;
SQ
Query Match 100.0%; Score 20; DB 22; Length 50;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AAPM 4
DB 31 AAPM 34
RESULT 35
AAU65105
ID AAU65105 standard; Protein; 50 AA.
XX
XX AAU65105;
AC
XX
XX 27-FEB-2002 (first entry)
DT
XX
XX Propionibacterium acnes immunogenic protein #26001.
DE
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
```


KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59659.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 26300; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 50 AA;
XX
Query Match 100.0%; Score 20; DB 22; Length 50;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPM 4
Db 46 AAPM 49
XX
RESULT 36
AAY65229
ID AAY65229 standard; Protein; 51 AA.
XX
AC AAY65229;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human 5' EST; related polypeptide SEQ ID NO:1390.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
KW gene therapy; chromosome mapping; upstream regulatory sequence;
XX

KW forensic; location; development; protein synthesis; stability;
KW regulation; identification.
XX
OS Homo sapiens.
XX
PN WO9953051-A2.
XX
PD 21-OCT-1999.
XX
PF 09-APR-1999; 99WO-IB00712.
XX
PR 09-APR-1998; 98US-0057719.
PR 28-APR-1998; 98US-0069047.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-038446/03.
DR N-PSDB; AA242843.
XX
XX Novel secreted protein 5' expressed sequence tag sequences used in
PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX
PS Claim 3; Page 771; 837pp; English.
XX
CC AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. AAY64651 to
CC AAY65438 represent the EST-related proteins corresponding to AA242265 to
CC AA243052. The 5' ESTs can be used for producing secreted human gene
CC products. They can be used to identify and isolate 5' untranslated
CC regions (UTRs) and upstream regulatory regions which control the
CC location, development stage, rate, and quantity of protein synthesis, as
CC well as stability of mRNA. The ESTs are also useful as probes for
CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
CC also be used in forensic procedures to identify individuals, or in
CC diagnostic procedures to identify individuals having genetic diseases
CC resulting from abnormal gene expression. The products may also be used in
CC gene therapy protocols. The nucleic acids encoding signal peptides can be
CC used for directing extracellular secretion of a polypeptide or the
CC insertion of a polypeptide into a membrane, or importing a polypeptide
CC into a cell. The proteins encoded by the EST sequences may be useful in
CC treating a variety of human conditions. Secreted proteins have
CC therapeutic value, and the identification of new secreted proteins is
CC valuable. AA242249 to AA242264 and AAY64644 to AAY64650 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 51 AA;
XX
Query Match 100.0%; Score 20; DB 21; Length 51;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPM 4
Db 42 AAPM 45
XX
RESULT 37
ABG00125
ID ABG00125 standard; Protein; 53 AA.
XX
AC ABG00125;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #116.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX

PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS64312.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID NO 30484; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 53 AA;
XX
Query Match 100.0%; Score 20; DB 22; Length 53;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AAPM 4
Db 21 AAPM 24
XX
RESULT 38
AAU64950
ID AU64950 standard; Protein; 56 AA.
XX
AC AAU64950;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #25846.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX

PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59655.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID NO 26145; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 56 AA;
XX
Query Match 100.0%; Score 20; DB 22; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AAPM 4
Db 17 AAPM 20
XX
RESULT 39
AAU67577
ID AU67577 standard; Protein; 59 AA.
XX
AC AAU67577;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #28473.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX

PN WO200181581-A2.
XX
PC 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59575.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID NO 28772; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 59 AA;

Query Match 100.0%; Score 20; DB 22; Length 59;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
DB 48 AAPM 51

RESULT 40
ABP10723
ID ABP10723 standard; Protein; 60 AA.
XX
AC ABP10723;
XX
DT 25-JUN-2002 (first entry)
XX
DE Human ORFX protein sequence SEQ ID NO:21428.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;

KW myasthenia gravis.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US10836.
XX
PR 30-MAY-2000; 2000US-206132P.
PR 29-AUG-2000; 2000US-228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach MD;
XX
DR WPI; 2002-106308/14.
DR N-PSDB; ABN26475.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders -
XX
PS Disclosure; SEQ ID 21428; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification)). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or peridontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 AA;

Query Match 100.0%; Score 20; DB 23; Length 60;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
DB 48 AAPM 51

Search completed: December 6, 2002, 13:28:01
Job time : 26 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:26:14 ; Search time 9.33333 Seconds
(without alignments)
12.610 Million cell updates/sec

Title: US-10-033-526-2
Perfect score: 20
Sequence: 1 AAPM 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 20 | 100.0 | 4 | 1 US-08-004-643C-3 | Sequence 3, Appl1 |
| 2 | 20 | 100.0 | 4 | 1 US-08-544-143A-4 | Sequence 4, Appl1 |
| 3 | 20 | 100.0 | 4 | 1 US-08-460-343B-60 | Sequence 60, Appl1 |
| 4 | 20 | 100.0 | 4 | 1 US-08-777-208-3 | Sequence 3, Appl1 |
| 5 | 20 | 100.0 | 4 | 1 US-08-398-028B-60 | Sequence 60, Appl1 |
| 6 | 20 | 100.0 | 4 | 1 US-08-693-653-2 | Sequence 2, Appl1 |
| 7 | 20 | 100.0 | 4 | 2 US-08-504-265B-60 | Sequence 60, Appl1 |
| 8 | 20 | 100.0 | 4 | 2 US-08-025-321C-13 | Sequence 13, Appl1 |
| 9 | 20 | 100.0 | 4 | 2 US-08-907-840A-1 | Sequence 1, Appl1 |
| 10 | 20 | 100.0 | 4 | 3 US-08-950-618-4 | Sequence 4, Appl1 |
| 11 | 20 | 100.0 | 4 | 3 US-08-950-618-19 | Sequence 19, Appl1 |
| 12 | 20 | 100.0 | 4 | 3 US-08-950-618-32 | Sequence 32, Appl1 |
| 13 | 20 | 100.0 | 4 | 3 US-08-950-618-33 | Sequence 33, Appl1 |
| 14 | 20 | 100.0 | 4 | 4 US-09-168-010-4 | Sequence 4, Appl1 |
| 15 | 20 | 100.0 | 4 | 4 US-09-168-010-19 | Sequence 19, Appl1 |
| 16 | 20 | 100.0 | 4 | 4 US-09-168-010-32 | Sequence 32, Appl1 |
| 17 | 20 | 100.0 | 4 | 4 US-09-168-010-33 | Sequence 33, Appl1 |
| 18 | 20 | 100.0 | 4 | 4 US-09-343-650-4 | Sequence 19, Appl1 |
| 19 | 20 | 100.0 | 4 | 4 US-09-343-650-19 | Sequence 19, Appl1 |
| 20 | 20 | 100.0 | 4 | 4 US-09-343-650-32 | Sequence 32, Appl1 |
| 21 | 20 | 100.0 | 4 | 4 US-09-343-650-33 | Sequence 33, Appl1 |
| 22 | 20 | 100.0 | 4 | 4 US-09-343-303-11 | Sequence 11, Appl1 |
| 23 | 20 | 100.0 | 6 | 4 US-08-988-842-5 | Sequence 5, Appl1 |
| 24 | 20 | 100.0 | 8 | 6 5194596-23 | Patent No. 5194596 |
| 25 | 20 | 100.0 | 8 | 6 5219739-28 | Patent No. 5219739 |
| 26 | 20 | 100.0 | 106 | 4 US-09-188-930-136 | Sequence 136, App |
| 27 | 20 | 100.0 | 110 | 4 US-08-905-223-336 | Sequence 336, App |

| | | | | | |
|----|----|-------|-----|-----------------------|--------------------|
| 28 | 20 | 100.0 | 116 | 4 US-09-134-001C-3725 | Sequence 3725, Ap |
| 29 | 20 | 100.0 | 123 | 3 US-08-840-316-3 | Sequence 3, Appl1 |
| 30 | 20 | 100.0 | 123 | 3 US-08-478-507-9 | Sequence 9, Appl1 |
| 31 | 20 | 100.0 | 123 | 4 US-08-809-523-3 | Sequence 3, Appl1 |
| 32 | 20 | 100.0 | 123 | 4 US-09-128-275A-9 | Sequence 9, Appl1 |
| 33 | 20 | 100.0 | 123 | 4 US-08-471-971-3 | Sequence 3, Appl1 |
| 34 | 20 | 100.0 | 123 | 4 US-09-553-427-9 | Sequence 9, Appl1 |
| 35 | 20 | 100.0 | 123 | 4 US-09-462-606-13 | Sequence 13, Appl1 |
| 36 | 20 | 100.0 | 123 | 4 US-09-462-606-58 | Sequence 58, Appl1 |
| 37 | 20 | 100.0 | 123 | 4 US-09-462-606-59 | Sequence 59, Appl1 |
| 38 | 20 | 100.0 | 123 | 4 US-09-462-606-60 | Sequence 60, Appl1 |
| 39 | 20 | 100.0 | 123 | 4 US-09-462-606-61 | Sequence 61, Appl1 |
| 40 | 20 | 100.0 | 123 | 4 US-09-462-606-62 | Sequence 62, Appl1 |
| 41 | 20 | 100.0 | 123 | 4 US-09-462-606-63 | Sequence 63, Appl1 |
| 42 | 20 | 100.0 | 123 | 4 US-09-462-606-64 | Sequence 64, Appl1 |
| 43 | 20 | 100.0 | 123 | 4 US-09-402-776-3 | Sequence 3, Appl1 |
| 44 | 20 | 100.0 | 123 | 5 PCT-US93-08849A-3 | Sequence 3, Appl1 |
| 45 | 20 | 100.0 | 123 | 5 PCT-US93-08849-3 | Sequence 3, Appl1 |

ALIGNMENTS

RESULT 1
US-08-004-643C-3
; Sequence 3, Application US/08004643C
; Patent No. 5480779
; GENERAL INFORMATION:
; APPLICANT: Gunter Fischer & Gerhard K Ilerz
; TITLE OF INVENTION: Cyclosporine Assay
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gunter Fischer et al. c/o G. P. Katona
; STREET: 230 Park Avenue, Room 2200
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10169
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: conf. to Patentln Release #1.0, Ver. #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/004, 643C
; FILING DATE: 12 January 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 703, 590
; FILING DATE: 20 May 1991
; APPLICATION NUMBER: 398, 092
; FILING DATE: 24 August 1989
; APPLICATION NUMBER: DD WP 601 F/319 577W
; FILING DATE: 07 September 1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Katona, Gabriel P.
; REGISTRATION NUMBER: 20, 829
; REFERENCE/DOCKET NUMBER: 691-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-3377
; TELEFAX: (212)986-6126
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-004-643C-3
Query Match 100.0%; Score 20; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 1 AAPM 4

RESULT 2

US-08-544-143A-4
; Sequence 4, Application US/08544143A
; Patent No. 5646028
; GENERAL INFORMATION:
; APPLICANT: Leigh, Scott D.
; TITLE OF INVENTION: NOVEL ALKALINE PROTEASE AND USES
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Majestic, Parsons, Siebert & Hsue
; STREET: Four Embarcadero Center, Suite 1100
; CITY: San Francisco
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94111-4121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/544,143A
; FILING DATE: 17-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: de Runtz, K. Alison
; REGISTRATION NUMBER: 37,119
; REFERENCE/DOCKET NUMBER: 0409.054US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-362-5556
; TELEFAX: 415-362-5418
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-544-143A-4

Query Match 100.0%; Score 20; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 1 AAPM 4

RESULT 3

US-08-460-343B-60
; Sequence 60, Application US/08460343B
; Patent No. 5741664
; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,343B
; FILING DATE: 01-Jun-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/398028
; FILING DATE: 03-mar-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: Amino Acid
; TOPOLOGY: linear
; US-08-460-343B-60

Query Match 100.0%; Score 20; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 1 AAPM 4

RESULT 4

US-08-777-208-3
; Sequence 3, Application US/08777208
; Patent No. 5763576
; GENERAL INFORMATION:
; APPLICANT: Powers, James C.
; TITLE OF INVENTION: Tetrapeptide Alpha-Ketoamides
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Deveau, Colton & Marquis
; STREET: Two Midtown Plaza, Suite 1400
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/777,208
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/539944
; FILING DATE: 06-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Colton, Laurence P.
; REGISTRATION NUMBER: 33371
; REFERENCE/DOCKET NUMBER: 10733-191B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 875-3555
; TELEFAX: (404) 875-8505
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
; HYPOTHETICAL: no
US-08-777-208-3

Query Match 100.0%; Score 20; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 1 AAPM 4

RESULT 5

US-08-398-028B-60
; Sequence 60, Application US/08398028B
; Patent No. 5780285

; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/398,028B
FILING DATE: 03-Mar-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kubinec, Jeffrey S.

REGISTRATION NUMBER: 36,575

REFERENCE/DOCKET NUMBER: P0936

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-8228

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 60:

SEQUENCE CHARACTERISTICS:

LENGTH: 4 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

US-08-398-028B-60

Query Match 100.0%; Score 20; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 1 AAPM 4

RESULT 6

US-08-693-653-2
; Sequence 2, Application US/08693653
; Patent No. 5780439

; GENERAL INFORMATION:
; APPLICANT: Mendy, Francois
; APPLICANT: Kahn, Jean-Maurice
; APPLICANT: Roger, Loic
; TITLE OF INVENTION: Improvements in or relating to organic
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Sandoz Corporation
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: NJ07936

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/693,653

FILING DATE: 09-AUG-1996

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/353,652

FILING DATE:

APPLICATION NUMBER: US 07/960,143

FILING DATE: 13-OCT-1992

ATTORNEY/AGENT INFORMATION:

NAME: Battle, Carl

REGISTRATION NUMBER: 30,731

REFERENCE/DOCKET NUMBER: 510-5747

TELECOMMUNICATION INFORMATION:

TELEPHONE: (201) 503-8532

TELEFAX: (201) 503-8807

TELEX: 240867

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 4 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Synthetic

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1

OTHER INFORMATION: /label= 2a

OTHER INFORMATION: /note= "succinyl derivative"

FEATURE:

NAME/KEY: Modified-site

LOCATION: 4

OTHER INFORMATION: /label= 2b

OTHER INFORMATION: /note= "-p-nitro-anilide derivative"

US-08-693-653-2

Query Match 100.0%; Score 20; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 1 AAPM 4

RESULT 7

US-08-504-265B-60
; Sequence 60, Application US/08504265B
; Patent No. 5837516

; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco

```
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,265B
; FILING DATE: 19-Jul-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/398028
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kuhnec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-08-504-265B-60
```

```
Query Match          100.0%; Score 20; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    1111
Db 1 AAPM 4
```

```
RESULT 8
US-08-025-321C-13
; Sequence 13, Application US/08025321C
; Patent No. 5849560
; GENERAL INFORMATION:
; APPLICANT: Abraham Ph.D., Carmela R.
; TITLE OF INVENTION: PROTEASES CAUSING ABNORMAL DEGRADATION
; TITLE OF INVENTION: OF AMYLOID BETA-PROTEIN PRECURSOR
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Choate, Hall & Stewart
; STREET: 53 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/025,321C
; FILING DATE: 26-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jarrell Ph.D., Brenda H.
; REGISTRATION NUMBER: 39,223
; REFERENCE/DOCKET NUMBER: 0079571-0034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 248 5000
; TELEFAX: 617 248 4000
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
```

```
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
;
US-08-025-321C-13
```

```
Query Match          100.0%; Score 20; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    1111
Db 1 AAPM 4
```

```
RESULT 9
US-08-907-840A-1
; Sequence 1, Application US/08907840A
; Patent No. 5952307
; GENERAL INFORMATION:
; APPLICANT: Powers, James C.
; TITLE OF INVENTION: Basic Alpha-Aminoalkylphosphonate
; TITLE OF INVENTION: Derivatives
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Deveau, Colton & Marquis
; STREET: Two Midtown Plaza, Suite 1400
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/907,840A
; FILING DATE: 14 AUG-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/184286
; FILING DATE: 21 JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Colton, Laurence P.
; REGISTRATION NUMBER: 33371
; REFERENCE/DOCKET NUMBER: 10733-175CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 875-3555
; TELEFAX: (404) 875-8505
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: no
;
US-08-907-840A-1
```

```
Query Match          100.0%; Score 20; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    1111
Db 1 AAPM 4
```

```
RESULT 10
US-08-950-618-4
; Sequence 4, Application US/08950618
; Patent No. 6087325
; GENERAL INFORMATION:
```

```
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: Ali, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; TITLE OF INVENTION: Methods of Liposomal Drug Delivery Using
; TITLE OF INVENTION: Peptidase-Cleavable Fusogenic Phospholipids
; FILE REFERENCE: TLC 215A
; CURRENT APPLICATION NUMBER: US/08/950,618
; EARLIER FILING DATE: 1997-10-15
; EARLIER APPLICATION NUMBER: US 60/027,544
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
US-08-950-618-4
```

```
Query Match          100.0%; Score 20; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPM 4
        ||||
Db       1 AAPM 4
```

```
RESULT 11
US-08-950-618-19
; Sequence 19, Application US/08950618
; Patent No. 6087325
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: Ali, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; TITLE OF INVENTION: Methods of Liposomal Drug Delivery Using
; TITLE OF INVENTION: Peptidase-Cleavable Fusogenic Phospholipids
; FILE REFERENCE: TLC 215A
; CURRENT APPLICATION NUMBER: US/08/950,618
; EARLIER FILING DATE: 1997-10-15
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-terminal methoxy succinyl group
US-08-950-618-19
```

```
Query Match          100.0%; Score 20; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPM 4
        ||||
Db       1 AAPM 4
```

```
RESULT 12
US-08-950-618-32
```

```
; Sequence 32, Application US/08950618
; Patent No. 6087325
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: Ali, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; TITLE OF INVENTION: Methods of Liposomal Drug Delivery Using
; TITLE OF INVENTION: Peptidase-Cleavable Fusogenic Phospholipids
; FILE REFERENCE: TLC 215A
; CURRENT APPLICATION NUMBER: US/08/950,618
; EARLIER FILING DATE: 1997-10-15
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-terminal succinyl group
US-08-950-618-32
```

```
Query Match          100.0%; Score 20; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPM 4
        ||||
Db       1 AAPM 4
```

```
RESULT 13
US-08-950-618-33
; Sequence 33, Application US/08950618
; Patent No. 6087325
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: Ali, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; TITLE OF INVENTION: Methods of Liposomal Drug Delivery Using
; TITLE OF INVENTION: Peptidase-Cleavable Fusogenic Phospholipids
; FILE REFERENCE: TLC 215A
; CURRENT APPLICATION NUMBER: US/08/950,618
; EARLIER FILING DATE: 1997-10-15
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-terminal carboxy sugar group
US-08-950-618-33
```

```
Query Match          100.0%; Score 20; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 AAPM 4
1111
Db 1 AAPM 4

RESULT 14

US-09-168-010-4
; Sequence 4, Application US/09168010A
; Patent No. 6143716
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: All, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; APPLICANT: Cabral-Lilly, Donna
; TITLE OF INVENTION: Liposomal Peptide-Lipid Conjugates and
; TITLE OF INVENTION: Delivery Using Same
; FILE REFERENCE: TLC 215B
; CURRENT APPLICATION NUMBER: US/09/168,010A
; CURRENT FILING DATE: 1998-10-07
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
US-09-168-010-4

Query Match 100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 1 AAPM 4

RESULT 15

US-09-168-010-19
; Sequence 19, Application US/09168010A
; Patent No. 6143716
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: All, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; APPLICANT: Cabral-Lilly, Donna
; TITLE OF INVENTION: Liposomal Peptide-Lipid Conjugates and
; TITLE OF INVENTION: Delivery Using Same
; FILE REFERENCE: TLC 215B
; CURRENT APPLICATION NUMBER: US/09/168,010A
; CURRENT FILING DATE: 1998-10-07
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-terminal methoxy succinyl group
US-09-168-010-19

Query Match 100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 1 AAPM 4

RESULT 16

US-09-168-010-32
; Sequence 32, Application US/09168010A
; Patent No. 6143716
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: All, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; APPLICANT: Cabral-Lilly, Donna
; TITLE OF INVENTION: Liposomal Peptide-Lipid Conjugates and
; TITLE OF INVENTION: Delivery Using Same
; FILE REFERENCE: TLC 215B
; CURRENT APPLICATION NUMBER: US/09/168,010A
; CURRENT FILING DATE: 1998-10-07
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-terminal succinyl group
US-09-168-010-32

Query Match 100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 1 AAPM 4

RESULT 17

US-09-168-010-33
; Sequence 33, Application US/09168010A
; Patent No. 6143716
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: All, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; APPLICANT: Cabral-Lilly, Donna
; TITLE OF INVENTION: Liposomal Peptide-Lipid Conjugates and
; TITLE OF INVENTION: Delivery Using Same
; FILE REFERENCE: TLC 215B
; CURRENT APPLICATION NUMBER: US/09/168,010A
; CURRENT FILING DATE: 1998-10-07
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSEQ for Windows Version 3.0

```
; SEQ ID NO 33
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-terminal carboxy sugar group
US-09-168-010-33

Query Match          100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
        ||||
Db       1 AAPM 4

RESULT 18
US-09-343-650-4
; Sequence 4, Application US/09343650
; Patent No. 6339069
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Meers, Paul
; APPLICANT: Pak, Charles
; APPLICANT: Ali, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; APPLICANT: Cabral-Lilly, Donna
; TITLE OF INVENTION: Peptide-Lipid Conjugates, Liposomes and
; FILE REFERENCE: TLC 215c
; CURRENT APPLICATION NUMBER: US/09/343,650
; CURRENT FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: US 08/950,618
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
US-09-343-650-4

Query Match          100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
        ||||
Db       1 AAPM 4

RESULT 19
US-09-343-650-19
; Sequence 19, Application US/09343650
; Patent No. 6339069
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Meers, Paul
; APPLICANT: Pak, Charles
; APPLICANT: Ali, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; APPLICANT: Cabral-Lilly, Donna
; TITLE OF INVENTION: Peptide-Lipid Conjugates, Liposomes and
; FILE REFERENCE: TLC 215c
; CURRENT APPLICATION NUMBER: US/09/343,650

; CURRENT FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: US 08/950,618
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-terminal succinyl group
US-09-343-650-32

Query Match          100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
        ||||
Db       1 AAPM 4

RESULT 20
US-09-343-650-32
; Sequence 32, Application US/09343650
; Patent No. 6339069
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Meers, Paul
; APPLICANT: Pak, Charles
; APPLICANT: Ali, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; APPLICANT: Cabral-Lilly, Donna
; TITLE OF INVENTION: Peptide-Lipid Conjugates, Liposomes and
; FILE REFERENCE: TLC 215c
; CURRENT APPLICATION NUMBER: US/09/343,650
; CURRENT FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: US 08/950,618
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-terminal succinyl group
US-09-343-650-32

Query Match          100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
        ||||
Db       1 AAPM 4

RESULT 21
US-09-343-650-33
; Sequence 33, Application US/09343650
; Patent No. 6339069
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Meers, Paul
; APPLICANT: Pak, Charles
; APPLICANT: Ali, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; APPLICANT: Cabral-Lilly, Donna
; TITLE OF INVENTION: Peptide-Lipid Conjugates, Liposomes and
; FILE REFERENCE: TLC 215c
; CURRENT APPLICATION NUMBER: US/09/343,650
```



```
; APPLICANT: Ali, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; APPLICANT: Cabral-Lilly, Donna
; TITLE OF INVENTION: Peptide-Lipid Conjugates, Liposomes and
; TITLE OF INVENTION: Liposomal Drug Delivery
; FILE REFERENCE: TLC 215C
; CURRENT APPLICATION NUMBER: US/09/343,650
; CURRENT FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: US 08/950,618
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-terminal carboxy sugar group
US-09-343-650-33
```

```
Query Match          100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    ||||
Db 1 AAPM 4
```

```
RESULT 22
US-09-578-303-11
; Sequence 11, Application US/09578303
; Patent No. 6399759
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Whitworth, S. Troy
; APPLICANT: Blum, Murray S.
; TITLE OF INVENTION: Ant Proteases and Methods of Inhibition
; FILE REFERENCE: 235.00150101
; CURRENT APPLICATION NUMBER: US/09/578,303
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/136,331
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: para-nitroanilide substrate
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: succinyl end cap
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: para-nitroanilide end cap
US-09-578-303-11
```

```
Query Match          100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    ||||
Db 1 AAPM 4
```

```
RESULT 23
US-08-988-842-5
; Sequence 5, Application US/08988842
; Patent No. 6462173
; GENERAL INFORMATION:
; APPLICANT: Lu, Jun Ping
; APPLICANT: Cantley, Lewis C.
; APPLICANT: Yaffe, Michael
; APPLICANT: Fischer, Gunter
; TITLE OF INVENTION: INHIBITORS OF PHOSPHOSERINE AND
; TITLE OF INVENTION: PHOSPHOTHEONINE-PROLINE-SPECIFIC ISOMERASES
; FILE REFERENCE: BIDMC97-02PA
; CURRENT APPLICATION NUMBER: US/08/988,842
; CURRENT FILING DATE: 1997-12-11
; EARLIER APPLICATION NUMBER: 60/058,164
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: synthetic nucleotide
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: (5)...(5)
US-08-988-842-5
```

```
Query Match          100.0%; Score 20; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    ||||
Db 1 AAPM 4
```

```
RESULT 24
5194596-23
; Patent No. 5194596
; APPLICANT: TISCHER, EDMUND G.;ABRAHAM, JUDITH A.;FIDDES, JOHN
; C.;MITCHELL, RICHARD L.
; TITLE OF INVENTION: PRODUCTION OF VASCULAR ENDOTHELIAL CELL
; GROWTH FACTOR
; NUMBER OF SEQUENCES: 32
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/450,883
; FILING DATE: 14-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 387,545
; FILING DATE: 27-JUL-1989
; SEQ ID NO:23:
; LENGTH: 8
5194596-23
```

```
Query Match          100.0%; Score 20; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    ||||
Db 4 AAPM 7
```

```
RESULT 25
5219739-28
; Patent No. 5219739
; APPLICANT: TISCHER, EDMUND G.;ABRAHAM, JUDITH A.;FIDDES,
; JOHN C.;MITCHELL, RICHARD L.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING BVGEF120 AND
; HVEGF 121 AND METHODS FOR THE PRODUCTION OF BOVINE AND HUMAN
; VASCULAR ENDOTHELIAL CELL GROWTH FACTORS, BVGEF120 AND HVEGF121
; NUMBER OF SEQUENCES: 40
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/07/559,041
; FILING DATE: 27-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 450, 883
; FILING DATE: 14-DEC-1989
; APPLICATION NUMBER: 387, 545
; FILING DATE: 27-JUL-1989
; SEQ ID NO: 28
; LENGTH: 8
5219739-28

Query Match          100.0%; Score 20; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
        ||||
Db       4 AAPM 7

RESULT 26
US-09-188-930-136
; Sequence 136, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Iorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 136
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Rat
US-09-188-930-136

Query Match          100.0%; Score 20; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
        ||||
Db       2 AAPM 5

RESULT 27
US-08-905-223-336
; Sequence 336, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Dueiert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
```

```
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelson, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 336:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -104...-1
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 5.1
; OTHER INFORMATION: seq SSVASLTATPSLA/SP
US-08-905-223-336

Query Match          100.0%; Score 20; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
        ||||
Db       71 AAPM 74

RESULT 28
US-09-134-001C-3725
; Sequence 3725, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3725
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3725

Query Match          100.0%; Score 20; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
        ||||
Db       29 AAPM 32

RESULT 29
US-08-840-316-3
; Sequence 3, Application US/08840316
; Patent No. 6054567
```

GENERAL INFORMATION:
APPLICANT: Emerson, Suzanne U., Purcell, Robert H.,
APPLICANT: Tsarev, Sergei. A., and Robinson, Robin A.
TITLE OF INVENTION: Recombinant Proteins Of
TITLE OF INVENTION: A Pakistani Strain Of Hepatitis E And Their
TITLE OF INVENTION: Use In Diagnostic Methods And Vaccines
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840,316
FILING DATE: 11-APR-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Richard W. Bork
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4255
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acid residues
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-840-316-3

Query Match 100.0%; Score 20; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 7 AAPM 10

RESULT 30
US-08-478-507-9
Sequence 9, Application US/08478507
Patent No. 6120988
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R
APPLICANT: Yarbough, Patrice O
APPLICANT: Bradley, Daniel W
APPLICANT: Krawczynski, Krzysztof Z
APPLICANT: Tam, Albert
APPLICANT: Fry, Kirk E
TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
TITLE OF INVENTION: NO. 6120988-A/No. 6120988-B Hepatitis Viral Agent
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,507
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/279,823
FILING DATE: 25-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681,078
FILING DATE: 05-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/336,672
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0183.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-507-9

Query Match 100.0%; Score 20; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 7 AAPM 10

RESULT 31
US-08-809-523-3
Sequence 3, Application US/08809523
Patent No. 6207416
GENERAL INFORMATION:
APPLICANT: Tsarev, Sergei. A., Emerson,
APPLICANT: Suzanne U., Purcell, Robert H.
TITLE OF INVENTION: Recombinant Proteins Of
TITLE OF INVENTION: A Pakistani Strain Of Hepatitis E And Their
TITLE OF INVENTION: Use In Diagnostic Methods And Vaccines
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,523
FILING DATE: 28-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13102
FILING DATE: 03-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US08/316,765
FILING DATE: 03-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/947,263
FILING DATE: 18-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Richard W. Bork
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4032US4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acid residues
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-809-523-3

Query Match 100.0%; Score 20; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
|||
Db 7 AAPM 10

RESULT 32
US-09-128-275A-9
Sequence 9, Application US/09128275A
Patent No. 6229005
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R
APPLICANT: Yarbough, Patrice O
APPLICANT: Bradley, Daniel W
APPLICANT: Krawczynski, Krzysztof Z
APPLICANT: Fry, Kirk E
TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/128,275A
FILING DATE: 03-AUG-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/279,823
FILING DATE: 25-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681,078
FILING DATE: 05-APR-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/336,672
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Petithory, Joanne R.
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 4600-0183.24
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-128-275A-9

Query Match 100.0%; Score 20; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
|||
Db 7 AAPM 10

RESULT 33
US-08-471-971-3
Sequence 3, Application US/08471971
Patent No. 6287759
GENERAL INFORMATION:
APPLICANT: Tsarev, Sergel A., Emerson,
APPLICANT: Suzanne U., Purcell, Robert H.
TITLE OF INVENTION: Recombinant Proteins Of
TITLE OF INVENTION: A Pakistani Strain Of Hepatitis E And Their
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,971
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US08/316,765
FILING DATE: 03-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US07/947,263
FILING DATE: 18-SEP-1992

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Richard W. Bork
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4032US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acid residues
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-471-971-3

Query Match 100.0%; Score 20; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 7 AAPM 10

RESULT 34
US-09-553-427-9
Sequence 9, Application US/09553427
Patent No. 6379891
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R
APPLICANT: Yarbough, Patrice O
APPLICANT: Bradley, Daniel W
APPLICANT: Krawczynski, Krzysztof Z
APPLICANT: Tam, Albert
APPLICANT: Fry, Kirk E
TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
TITLE OF INVENTION: No. 6379891-A/No. 6379891-B Hepatitis Viral Agent
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/553,427
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/478,507
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/279,823
FILING DATE: 25-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681,078
FILING DATE: 05-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/336,672

FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0183.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-553-427-9

Query Match 100.0%; Score 20; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 7 AAPM 10

RESULT 35
US-09-462-606-13
Sequence 13, Application US/09462606
Patent No. 6432408
GENERAL INFORMATION:
APPLICANT: MENG, XIANG-JIN
APPLICANT: Emerson, Suzanne U.
APPLICANT: Purcell, Robert H.
TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
FILE REFERENCE: 20264267US1
CURRENT APPLICATION NUMBER: US/09/462,606
CURRENT FILING DATE: 2000-06-12
PRIOR APPLICATION NUMBER: US 60/053069
PRIOR FILING DATE: 1997-07-18
PRIOR APPLICATION NUMBER: PCF/US98/14665
PRIOR FILING DATE: 1998-07-17
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 123
TYPE: PRT
ORGANISM: Hepatitis E virus
US-09-462-606-13

Query Match 100.0%; Score 20; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 7 AAPM 10

RESULT 36
US-09-462-606-58
Sequence 58, Application US/09462606
Patent No. 6432408
GENERAL INFORMATION:
APPLICANT: MENG, XIANG-JIN
APPLICANT: Emerson, Suzanne U.
APPLICANT: Purcell, Robert H.
TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
FILE REFERENCE: 20264267US1
CURRENT APPLICATION NUMBER: US/09/462,606
CURRENT FILING DATE: 2000-06-12


```
; PRIOR APPLICATION NUMBER: US 60/053069
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/14665
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 58
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-09-462-606-58

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 123;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 7 AAPM 10

RESULT 37
US-09-462-606-59
; Sequence 59, Application US/09462606
; Patent No. 6432408
; GENERAL INFORMATION:
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
; FILE REFERENCE: 20264267US1
; CURRENT APPLICATION NUMBER: US/09/462,606
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 60/053069
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/14665
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 59
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-09-462-606-59

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 123;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 7 AAPM 10

RESULT 38
US-09-462-606-60
; Sequence 60, Application US/09462606
; Patent No. 6432408
; GENERAL INFORMATION:
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
; FILE REFERENCE: 20264267US1
; CURRENT APPLICATION NUMBER: US/09/462,606
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 60/053069
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/14665
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 60
```

```
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-09-462-606-60

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 123;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 7 AAPM 10

RESULT 39
US-09-462-606-61
; Sequence 61, Application US/09462606
; Patent No. 6432408
; GENERAL INFORMATION:
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
; FILE REFERENCE: 20264267US1
; CURRENT APPLICATION NUMBER: US/09/462,606
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 60/053069
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/14665
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 61
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-09-462-606-61

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 123;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 7 AAPM 10

RESULT 40
US-09-462-606-62
; Sequence 62, Application US/09462606
; Patent No. 6432408
; GENERAL INFORMATION:
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
; FILE REFERENCE: 20264267US1
; CURRENT APPLICATION NUMBER: US/09/462,606
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 60/053069
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/14665
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 62
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-09-462-606-62

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 123;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

| | | | | | | | | | |
|---------|----|--------------|----|------------|----|--------|----|------|----|
| Matches | 4; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
| QY | 1 | AAPM | 4 | | | | | | |
| | | | | | | | | | |
| Db | 7 | AAPM | 10 | | | | | | |

Search completed: December 6, 2002, 13:31:50
Job time : 9.33333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:30:14 ; Search time 4.88889 Seconds
(without alignments)
13.289 Million cell updates/sec

Title: US-10-033-526-2
Perfect-score: 20
Sequence: 1 AAPM 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA: *
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | length | DB ID | Description |
|------------|-------|-------------|--------|------------------------|--------------------|
| 1 | 20 | 100.0 | 4 | 12 US-10-033-526-2 | Sequence 2, Appli |
| 2 | 20 | 100.0 | 31 | 10 US-09-864-761-48187 | Sequence 48187, A |
| 3 | 20 | 100.0 | 85 | 10 US-09-867-550-500 | Sequence 500, App |
| 4 | 20 | 100.0 | 112 | 10 US-09-764-869-939 | Sequence 939, App |
| 5 | 20 | 100.0 | 117 | 10 US-09-764-864-1182 | Sequence 1182, App |
| 6 | 20 | 100.0 | 120 | 10 US-09-917-265-21 | Sequence 21, Appli |
| 7 | 20 | 100.0 | 124 | 10 US-09-864-761-44327 | Sequence 44327, A |
| 8 | 20 | 100.0 | 124 | 10 US-09-769-066-21 | Sequence 21, Appli |
| 9 | 20 | 100.0 | 124 | 10 US-09-769-066-22 | Sequence 22, Appli |
| 10 | 20 | 100.0 | 126 | 10 US-09-815-242-4902 | Sequence 4902, App |
| 11 | 20 | 100.0 | 127 | 10 US-09-815-242-10542 | Sequence 10542, A |
| 12 | 20 | 100.0 | 128 | 10 US-09-864-761-43702 | Sequence 43702, A |
| 13 | 20 | 100.0 | 128 | 10 US-09-815-242-13224 | Sequence 13224, A |
| 14 | 20 | 100.0 | 132 | 10 US-09-864-761-36620 | Sequence 36620, A |
| 15 | 20 | 100.0 | 156 | 10 US-09-815-242-10329 | Sequence 10329, A |
| 16 | 20 | 100.0 | 156 | 10 US-09-815-242-13886 | Sequence 13886, A |
| 17 | 20 | 100.0 | 171 | 10 US-09-812-133-2 | Sequence 2, Appli |
| 18 | 20 | 100.0 | 191 | 9 US-09-870-759-122 | Sequence 122, App |
| 19 | 20 | 100.0 | 191 | 10 US-09-349-954A-2 | Sequence 2, Appli |

| 20 | 20 | 100.0 | 191 | 10 | US-09-932-451A-2 | Sequence 2, Appli |
|----|----|-------|-----|----|--------------------|--------------------|
| 21 | 20 | 100.0 | 191 | 10 | US-09-907-007-2 | Sequence 2, Appli |
| 22 | 20 | 100.0 | 191 | 10 | US-09-795-006A-2 | Sequence 2, Appli |
| 23 | 20 | 100.0 | 192 | 9 | US-09-852-209A-8 | Sequence 8, Appli |
| 24 | 20 | 100.0 | 198 | 9 | US-09-992-598-401 | Sequence 401, App |
| 25 | 20 | 100.0 | 198 | 9 | US-09-989-293A-401 | Sequence 401, App |
| 26 | 20 | 100.0 | 198 | 10 | US-09-989-722-401 | Sequence 401, App |
| 27 | 20 | 100.0 | 198 | 10 | US-09-989-723-401 | Sequence 401, App |
| 28 | 20 | 100.0 | 198 | 10 | US-09-989-279-401 | Sequence 401, App |
| 29 | 20 | 100.0 | 198 | 10 | US-09-989-727-401 | Sequence 401, App |
| 30 | 20 | 100.0 | 198 | 10 | US-09-989-731-401 | Sequence 401, App |
| 31 | 20 | 100.0 | 198 | 10 | US-09-989-732-401 | Sequence 401, App |
| 32 | 20 | 100.0 | 198 | 10 | US-09-991-073-401 | Sequence 401, App |
| 33 | 20 | 100.0 | 198 | 10 | US-09-990-442-401 | Sequence 401, App |
| 34 | 20 | 100.0 | 198 | 10 | US-09-991-163-401 | Sequence 401, App |
| 35 | 20 | 100.0 | 198 | 10 | US-09-993-604-401 | Sequence 401, App |
| 36 | 20 | 100.0 | 198 | 10 | US-09-990-456-401 | Sequence 401, App |
| 37 | 20 | 100.0 | 198 | 10 | US-09-989-721-401 | Sequence 401, App |
| 38 | 20 | 100.0 | 215 | 10 | US-09-244-694-3 | Sequence 3, Appli |
| 39 | 20 | 100.0 | 216 | 10 | US-09-291-299A-8 | Sequence 8, Appli |
| 40 | 20 | 100.0 | 221 | 10 | US-09-925-299-967 | Sequence 967, App |
| 41 | 20 | 100.0 | 228 | 9 | US-09-996-634-128 | Sequence 128, App |
| 42 | 20 | 100.0 | 232 | 10 | US-09-795-006A-147 | Sequence 147, App |
| 43 | 20 | 100.0 | 232 | 12 | US-10-127-551-5 | Sequence 5, Appli |
| 44 | 20 | 100.0 | 247 | 8 | US-08-450-842-3 | Sequence 3, Appli |
| 45 | 20 | 100.0 | 258 | 10 | US-09-810-264-32 | Sequence 32, Appli |

ALIGNMENTS

RESULT 1
US-10-033-526-2
; Sequence 2, Application US/10033526
; Patent No. US20020147999A1
; GENERAL INFORMATION:
; APPLICANT: Robert W. Mahley
; APPLICANT: Yadong Huang
; TITLE OF INVENTION: Methods of Treating Disorders Related to
; FILE OF INVENTION: APOE
; FILE REFERENCE: UCAL217
; CURRENT APPLICATION NUMBER: US/10/033.526
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 60/245,737
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-033-526-2

Query Match 100.0%; Score 20; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 1 AAPM 4

RESULT 2
US-09-864-761-48187
; Sequence 48187, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng

```

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48187
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL021707.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
US-09-864-761-48187

Query Match          100.0%; Score 20; DB 10; Length 31;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
      1111
Db      14 AAPM 17

RESULT 3
US-09-867-550-500
; Sequence 500, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James

```

```

; TITLE OF INVENTION: No. US20020082206A1e1 Polynucleotides from Atherogenic Cells a
; FILE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 500
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (69)
; OTHER INFORMATION: wherein xaa may be any one of Ile or Leu or Phe or Val
US-09-867-550-500

Query Match          100.0%; Score 20; DB 10; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
      1111
Db      44 AAPM 47

RESULT 4
US-09-764-869-939
; Sequence 939, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 939
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (63)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
US-09-764-869-939

Query Match          100.0%; Score 20; DB 10; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
      1111
Db      3 AAPM 6

RESULT 5
US-09-764-864-1182
; Sequence 1182, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17

```

```
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1182
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (93)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (116)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1182

Query Match          100.0%; Score 20; DB 10; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
        ||||
Db       34 AAPM 37

RESULT 6
US-09-917-265-21
; Sequence 21, Application US/09917265
; Patent No. US20020052030A1
; GENERAL INFORMATION:
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE PROTEINS, NUCLEIC ACID MOLECULES AND USES THERE
; FILE REFERENCE: IM-5
; CURRENT APPLICATION NUMBER: US/09/917,265
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/223,016
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 21
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Felis catus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (473)..(473)
; OTHER INFORMATION: n = unknown at position 473
US-09-917-265-21

Query Match          100.0%; Score 20; DB 10; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
        ||||
Db       9 AAPM 12

RESULT 7
US-09-864-761-44327
; Sequence 44327, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44327
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000134.14
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.57
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.79
; OTHER INFORMATION: EST_HUMAN HIT: AF001543.1, EVALUE 5.00e-41
; OTHER INFORMATION: SWISSPROT HIT: O14578, EVALUE 8.00e+00
US-09-864-761-44327

Query Match          100.0%; Score 20; DB 10; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
        ||||
Db      119 AAPM 122

RESULT 8
US-09-769-066-21
; Sequence 21, Application US/09769066
; Patent No. US20020107360A1
```


GENERAL INFORMATION:
APPLICANT: Fuerst, Thomas R.
McAtee, C. Patrick
Yarbough, Patrice O.
Zhang, Yifan
TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/769,066
FILING DATE: 24-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/542,634
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 124 amino acids
TYPE: amino acid
STRANDEDNESS: Hepatitis E Virus (Burma Strain)
ORF-3
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-769-066-21
Query Match 100.0%; Score 20; DB 10; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPM 4
1111
Db 7 AAPM 10
RESULT 9
US-09-769-066-22
Sequence 22, Application US/09769066
Patent No. US20020107360A1
GENERAL INFORMATION:
APPLICANT: Fuerst, Thomas R.
McAtee, C. Patrick
Yarbough, Patrice O.
Zhang, Yifan
TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/769,066
FILING DATE: 24-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/542,634
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 124 amino acids
TYPE: amino acid
STRANDEDNESS: Hepatitis E Virus (Mexico Strain)
ORF-3
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-769-066-22
Query Match 100.0%; Score 20; DB 10; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPM 4
1111
Db 7 AAPM 10
RESULT 10
US-09-815-242-4902
Sequence 4902, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4902
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-4902

Query Match 100.0%; Score 20; DB 10; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 117 AAPM 120

RESULT 11
US-09-815-242-10542
; Sequence 10542, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10542
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10542

Query Match 100.0%; Score 20; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 117 AAPM 120

RESULT 12
US-09-864-761-43702
; Sequence 43702, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43702
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003693.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EST_HUMAN HIT: AA419611.1, EVALUE 1.90e+00
US-09-864-761-43702

Query Match 100.0%; Score 20; DB 10; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 123 AAPM 126

RESULT 13
US-09-815-242-13224
; Sequence 13224, Application US/09815242
; Patent No. US20020061569A1

```
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13224
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-815-242-13224

Query Match      100.0%; Score 20; DB 10; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
Db      119 AAPM 122

RESULT 14
US-09-864-761-36620
; Sequence 36620, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36620
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL096774.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EST_HUMAN HIT: A1354615.1, EVALUE 1.60e-01
; OTHER INFORMATION: SWISSPROT HIT: O15945, EVALUE 6.50e+00
; US-09-864-761-36620

Query Match      100.0%; Score 20; DB 10; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
Db      89 AAPM 92

RESULT 15
US-09-815-242-10329
; Sequence 10329, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
```

; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10329
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10329

Query Match 100.0%; Score 20; DB 10; Length 156;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 49 AAPM 52

RESULT 16
US-09-815-242-13886
; Sequence 13886, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13886
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-13886

Query Match 100.0%; Score 20; DB 10; Length 156;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 49 AAPM 52

RESULT 17
US-09-812-133-2
; Sequence 2, Application US/09812133
; Patent No. US20020065240A1
; GENERAL INFORMATION:
; APPLICANT: Thomas, Kenneth A., Jr.
; APPLICANT: Bett, Andrew J.
; APPLICANT: Huckie, William R.
; TITLE OF INVENTION: GENE THERAPY FOR STIMULATION OF
; FILE REFERENCE: 20073P
; CURRENT APPLICATION NUMBER: US/09/812,133
; CURRENT FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: PCT/US98/22668
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/063,629
; PRIOR FILING DATE: 1997-10-27
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Human
US-09-812-133-2

Query Match 100.0%; Score 20; DB 10; Length 171;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 26 AAPM 29

RESULT 18
US-09-870-759-122
; Sequence 122, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 122
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-122

Query Match 100.0%; Score 20; DB 9; Length 191;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 26 AAPM 29

RESULT 19
US-09-349-954A-2
; Sequence 2, Application US/09349954A
; Patent No. US20020019027A1

```
; GENERAL INFORMATION:
; APPLICANT: Hayward, Nicholas K.
; APPLICANT: Weber, Gunther
; APPLICANT: Grimmold, Sean
; APPLICANT: NO. US20020019027AIdenskjold, Magnus
; APPLICANT: Larsson, Catharina
; TITLE OF INVENTION: A NOVEL GROWTH FACTOR AND A GENETIC SEQUENCE ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: Dav. Col. Cave
; CURRENT APPLICATION NUMBER: US/09/349,954A
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 08/765,588
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Nucleotide Sequence of VEGF165
US-09-349-954A-2
```

```
Query Match          100.0%; Score 20; DB 10; Length 191;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAPM 4
        ||||
Db       26 AAPM 29
```

```
RESULT 20
US-09-932-451A-2
; Sequence 2, Application US/09932451A
; Patent No. US20020111324A1
; GENERAL INFORMATION:
; APPLICANT: OZAWA, Kei-ya
; APPLICANT: SHIMPO, Masahisa
; APPLICANT: IKEDA, Uichi
; APPLICANT: MAEDA, Yoshikazu
; APPLICANT: SHIMADA, Kazuyuki
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS-MEDIATED DELIVERY OF ANGIOGENIC
; TITLE OF INVENTION: FACTORS
; FILE REFERENCE: 0800-0026
; CURRENT APPLICATION NUMBER: US/09/932,451A
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/226,056
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VEGF-165
US-09-932-451A-2
```

```
Query Match          100.0%; Score 20; DB 10; Length 191;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAPM 4
        ||||
Db       26 AAPM 29
```

```
RESULT 21
US-09-907-007-2
; Sequence 2, Application US/09907007
; Patent No. US20020142395A1
; GENERAL INFORMATION:
; APPLICANT: Hayward, Nicholas K.
; APPLICANT: Weber, Gunther
```

```
; APPLICANT: Grimmold, Sean
; APPLICANT: NO. US20020142395AIdenskjold, Magnus
; APPLICANT: Larsson, Catharina
; TITLE OF INVENTION: A NOVEL GROWTH FACTOR AND A GENETIC SEQUENCE ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: DAVIES
; CURRENT APPLICATION NUMBER: US/09/907,007
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 08/765,588
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Nucleotide Sequence of VEGF165
US-09-907-007-2
```

```
Query Match          100.0%; Score 20; DB 10; Length 191;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAPM 4
        ||||
Db       26 AAPM 29
```

```
RESULT 22
US-09-795-006A-2
; Sequence 2, Application US/09795006A
; Patent No. US20020151680A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al
; TITLE OF INVENTION: MATERIALS AND METHODS INVOLVING HYBRID VASCULAR
; TITLE OF INVENTION: ENDOTHELIAL GROWTH FACTOR DNAs AND PROTEINS
; FILE REFERENCE: 28967/35977B
; CURRENT APPLICATION NUMBER: US/09/795,006A
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: US 60/205,331
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/185,205
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 175
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-795-006A-2
```

```
Query Match          100.0%; Score 20; DB 10; Length 191;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAPM 4
        ||||
Db       26 AAPM 29
```

```
RESULT 23
US-09-852-209A-8
; Sequence 8, Application US/09852209A
; Patent No. US20020164687A1
; GENERAL INFORMATION:
; APPLICANT: ERIKSSON, Ulf
; APPLICANT: AASE, Karin
; APPLICANT: LEE, Xuri
; APPLICANT: PONTEN, Annica
; APPLICANT: UUTELA, Marko
; APPLICANT: ALITALO, Kari
; APPLICANT: OESTMAN, Arne
; APPLICANT: HELDIN, Carl-Henrik
; APPLICANT: BETSHOLTZ, Christer
```



```

; TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR C,      DNA CODING
; FILE REFERENCE: 09-410349-Eriksson et al-1064-44740
; CURRENT APPLICATION NUMBER: US/09/852,209A
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/410,349
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/110,749
; PRIOR FILING DATE: 1998-12-03
; PRIOR APPLICATION NUMBER: 60/113,002
; PRIOR FILING DATE: 1998-12-18
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 60/135,426
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-852-209A-8

Query Match          100.0%; Score 20; DB 9; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
   1111
Db 26 AAPM 29

RESULT 24
US-09-992-598-401
; Sequence 401, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC20
; CURRENT APPLICATION NUMBER: US/09/992,598
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311

```

```

; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16

```

;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478

;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 20; DB 9; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPM 4
Db 21 AAPM 24

RESULT 25
US-09-989-293A-401
; Sequence 401, Application US/09989293A
; Patent No. US20020177164A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C66
; CURRENT APPLICATION NUMBER: US/09/989,293A
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20

[illegible]

```
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      100.0%; Score 20; DB 9; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
       1111
Db      21 AAPM 24

RESULT 26
US-09-989-722-401
; Sequence 401, Application US/09989722
; Patent No. US20020072067A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C63
; CURRENT APPLICATION NUMBER: US/09/989,722
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
```

```
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
```

; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182

; PRIOR FILING DATE: 1998-07-09
Query Match 100.0%; Score 20; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPM 4
Db 21 AAPM 24
RESULT 27
US-09-989-723-401
; Sequence 401, Application US/09989723
; Patent No. US20020072092A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC62
; CURRENT APPLICATION NUMBER: US/09/989,723
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03

; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948

; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 20; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 21 AAPM 24

RESULT 28
US-09-989-279-401
; Sequence 401, Application US/09989279
; Patent No. US20020072496A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Guiney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C56
; CURRENT APPLICATION NUMBER: US/09/989,279
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028

; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22

; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 20; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 21 AAPM 24

RESULT 29
US-09-989-727-401
; Sequence 401, Application US/09989727

; Patent No. US20020072497A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C65
; CURRENT APPLICATION NUMBER: US/09/989,727
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 20; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 21 AAPM 24

RESULT 30
US-09-989-731-401
; Sequence 401, Application US/09989731
; Patent No. US20020103125A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C70
CURRENT APPLICATION NUMBER: US/09/989,731
PRIOR FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212

PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24


```

; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          100.0%; Score 20; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
       ||||
Db      21 AAPM 24

RESULT 31
US-09-989-732-401
; Sequence 401, Application US/09989732
; Patent No. US20020123463A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C57
; CURRENT APPLICATION NUMBER: US/09/989,732
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
```

```

; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540

```

```

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

```

```

Query Match          100.0%; Score 20; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 AAPM 4
      1111
Db      21 AAPM 24

```

```

RESULT 32
US-09-991-073-401
; Sequence 401, Application US/09991073
; Patent No. US20020127576A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.

```

APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC15
CURRENT APPLICATION NUMBER: US/09/991,073
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824

PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25

; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 20; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 21 AAPM 24

RESULT 33
US-09-990-442-401
; Sequence 401, Application US/09990442
; Patent No. US20020132252A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kjavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2730P1C8
; CURRENT APPLICATION NUMBER: US/09/990,442
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11

; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695

; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 20; DB 10; Length 198;
Best local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 21 AAPM 24

RESULT 34
US-09-991-163-401
; Sequence 401, Application US/09991163
; Patent No. US20020132253A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C17
; CURRENT APPLICATION NUMBER: US/09/991,163
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17

| | |
|----|-------------------------------------|
| 1 | PRIOR APPLICATION NUMBER: 60/065186 |
| 2 | PRIOR FILING DATE: 1997-11-12 |
| 3 | PRIOR APPLICATION NUMBER: 60/065311 |
| 4 | PRIOR FILING DATE: 1997-11-13 |
| 5 | PRIOR APPLICATION NUMBER: 60/066770 |
| 6 | PRIOR FILING DATE: 1997-11-24 |
| 7 | PRIOR APPLICATION NUMBER: 60/075945 |
| 8 | PRIOR FILING DATE: 1998-02-25 |
| 9 | PRIOR APPLICATION NUMBER: 60/078910 |
| 10 | PRIOR FILING DATE: 1998-03-20 |
| 11 | PRIOR APPLICATION NUMBER: 60/083322 |
| 12 | PRIOR FILING DATE: 1998-04-28 |
| 13 | PRIOR APPLICATION NUMBER: 60/084600 |
| 14 | PRIOR FILING DATE: 1998-05-07 |
| 15 | PRIOR APPLICATION NUMBER: 60/087106 |
| 16 | PRIOR FILING DATE: 1998-05-28 |
| 17 | PRIOR APPLICATION NUMBER: 60/087607 |
| 18 | PRIOR FILING DATE: 1998-06-02 |
| 19 | PRIOR APPLICATION NUMBER: 60/087609 |
| 20 | PRIOR FILING DATE: 1998-06-02 |
| 21 | PRIOR APPLICATION NUMBER: 60/087759 |
| 22 | PRIOR FILING DATE: 1998-06-02 |
| 23 | PRIOR APPLICATION NUMBER: 60/087827 |
| 24 | PRIOR FILING DATE: 1998-06-03 |
| 25 | PRIOR APPLICATION NUMBER: 60/088021 |
| 26 | PRIOR FILING DATE: 1998-06-04 |
| 27 | PRIOR APPLICATION NUMBER: 60/088025 |
| 28 | PRIOR FILING DATE: 1998-06-04 |
| 29 | PRIOR APPLICATION NUMBER: 60/088026 |
| 30 | PRIOR FILING DATE: 1998-06-04 |
| 31 | PRIOR APPLICATION NUMBER: 60/088028 |
| 32 | PRIOR FILING DATE: 1998-06-04 |
| 33 | PRIOR APPLICATION NUMBER: 60/088029 |
| 34 | PRIOR FILING DATE: 1998-06-04 |
| 35 | PRIOR APPLICATION NUMBER: 60/088030 |
| 36 | PRIOR FILING DATE: 1998-06-04 |
| 37 | PRIOR APPLICATION NUMBER: 60/088033 |
| 38 | PRIOR FILING DATE: 1998-06-04 |
| 39 | PRIOR APPLICATION NUMBER: 60/088326 |
| 40 | PRIOR FILING DATE: 1998-06-04 |
| 41 | PRIOR APPLICATION NUMBER: 60/088167 |
| 42 | PRIOR FILING DATE: 1998-06-05 |
| 43 | PRIOR APPLICATION NUMBER: 60/088202 |
| 44 | PRIOR FILING DATE: 1998-06-05 |
| 45 | PRIOR APPLICATION NUMBER: 60/088212 |
| 46 | PRIOR FILING DATE: 1998-06-05 |
| 47 | PRIOR APPLICATION NUMBER: 60/088217 |
| 48 | PRIOR FILING DATE: 1998-06-05 |
| 49 | PRIOR APPLICATION NUMBER: 60/088655 |
| 50 | PRIOR FILING DATE: 1998-06-09 |
| 51 | PRIOR APPLICATION NUMBER: 60/088734 |
| 52 | PRIOR FILING DATE: 1998-06-10 |
| 53 | PRIOR APPLICATION NUMBER: 60/088738 |
| 54 | PRIOR FILING DATE: 1998-06-10 |
| 55 | PRIOR APPLICATION NUMBER: 60/088742 |
| 56 | PRIOR FILING DATE: 1998-06-10 |
| 57 | PRIOR APPLICATION NUMBER: 60/088810 |
| 58 | PRIOR FILING DATE: 1998-06-10 |
| 59 | PRIOR APPLICATION NUMBER: 60/088824 |
| 60 | PRIOR FILING DATE: 1998-06-10 |
| 61 | PRIOR APPLICATION NUMBER: 60/088826 |
| 62 | PRIOR FILING DATE: 1998-06-10 |
| 63 | PRIOR APPLICATION NUMBER: 60/088858 |
| 64 | PRIOR FILING DATE: 1998-06-11 |
| 65 | PRIOR APPLICATION NUMBER: 60/088876 |
| 66 | PRIOR FILING DATE: 1998-06-11 |
| 67 | PRIOR APPLICATION NUMBER: 60/089105 |
| 68 | PRIOR FILING DATE: 1998-06-12 |
| 69 | PRIOR APPLICATION NUMBER: 60/089440 |
| 70 | PRIOR FILING DATE: 1998-06-16 |
| 71 | PRIOR APPLICATION NUMBER: 60/089512 |

| | |
|----|-------------------------------------|
| 1 | PRIOR FILING DATE: 1998-06-16 |
| 2 | PRIOR APPLICATION NUMBER: 60/089514 |
| 3 | PRIOR FILING DATE: 1998-06-16 |
| 4 | PRIOR APPLICATION NUMBER: 60/089532 |
| 5 | PRIOR FILING DATE: 1998-06-17 |
| 6 | PRIOR APPLICATION NUMBER: 60/089538 |
| 7 | PRIOR FILING DATE: 1998-06-17 |
| 8 | PRIOR APPLICATION NUMBER: 60/089598 |
| 9 | PRIOR FILING DATE: 1998-06-17 |
| 10 | PRIOR APPLICATION NUMBER: 60/089599 |
| 11 | PRIOR FILING DATE: 1998-06-17 |
| 12 | PRIOR APPLICATION NUMBER: 60/089600 |
| 13 | PRIOR FILING DATE: 1998-06-17 |
| 14 | PRIOR APPLICATION NUMBER: 60/089653 |
| 15 | PRIOR FILING DATE: 1998-06-17 |
| 16 | PRIOR APPLICATION NUMBER: 60/089801 |
| 17 | PRIOR FILING DATE: 1998-06-18 |
| 18 | PRIOR APPLICATION NUMBER: 60/089907 |
| 19 | PRIOR FILING DATE: 1998-06-18 |
| 20 | PRIOR APPLICATION NUMBER: 60/089908 |
| 21 | PRIOR FILING DATE: 1998-06-18 |
| 22 | PRIOR APPLICATION NUMBER: 60/089947 |
| 23 | PRIOR FILING DATE: 1998-06-19 |
| 24 | PRIOR APPLICATION NUMBER: 60/089948 |
| 25 | PRIOR FILING DATE: 1998-06-19 |
| 26 | PRIOR APPLICATION NUMBER: 60/089952 |
| 27 | PRIOR FILING DATE: 1998-06-19 |
| 28 | PRIOR APPLICATION NUMBER: 60/090246 |
| 29 | PRIOR FILING DATE: 1998-06-22 |
| 30 | PRIOR APPLICATION NUMBER: 60/090252 |
| 31 | PRIOR FILING DATE: 1998-06-22 |
| 32 | PRIOR APPLICATION NUMBER: 60/090254 |
| 33 | PRIOR FILING DATE: 1998-06-22 |
| 34 | PRIOR APPLICATION NUMBER: 60/090349 |
| 35 | PRIOR FILING DATE: 1998-06-23 |
| 36 | PRIOR APPLICATION NUMBER: 60/090355 |
| 37 | PRIOR FILING DATE: 1998-06-23 |
| 38 | PRIOR APPLICATION NUMBER: 60/090429 |
| 39 | PRIOR FILING DATE: 1998-06-24 |
| 40 | PRIOR APPLICATION NUMBER: 60/090431 |
| 41 | PRIOR FILING DATE: 1998-06-24 |
| 42 | PRIOR APPLICATION NUMBER: 60/090435 |
| 43 | PRIOR FILING DATE: 1998-06-24 |
| 44 | PRIOR APPLICATION NUMBER: 60/090444 |
| 45 | PRIOR FILING DATE: 1998-06-24 |
| 46 | PRIOR APPLICATION NUMBER: 60/090445 |
| 47 | PRIOR FILING DATE: 1998-06-24 |
| 48 | PRIOR APPLICATION NUMBER: 60/090472 |
| 49 | PRIOR FILING DATE: 1998-06-24 |
| 50 | PRIOR APPLICATION NUMBER: 60/090535 |
| 51 | PRIOR FILING DATE: 1998-06-24 |
| 52 | PRIOR APPLICATION NUMBER: 60/090540 |
| 53 | PRIOR FILING DATE: 1998-06-24 |
| 54 | PRIOR APPLICATION NUMBER: 60/090542 |
| 55 | PRIOR FILING DATE: 1998-06-24 |
| 56 | PRIOR APPLICATION NUMBER: 60/090557 |
| 57 | PRIOR FILING DATE: 1998-06-24 |
| 58 | PRIOR APPLICATION NUMBER: 60/090676 |
| 59 | PRIOR FILING DATE: 1998-06-25 |
| 60 | PRIOR APPLICATION NUMBER: 60/090678 |
| 61 | PRIOR FILING DATE: 1998-06-25 |
| 62 | PRIOR APPLICATION NUMBER: 60/090690 |
| 63 | PRIOR FILING DATE: 1998-06-25 |
| 64 | PRIOR APPLICATION NUMBER: 60/090694 |
| 65 | PRIOR FILING DATE: 1998-06-25 |
| 66 | PRIOR APPLICATION NUMBER: 60/090695 |
| 67 | PRIOR FILING DATE: 1998-06-25 |
| 68 | PRIOR APPLICATION NUMBER: 60/090696 |
| 69 | PRIOR FILING DATE: 1998-06-25 |
| 70 | PRIOR APPLICATION NUMBER: 60/090862 |
| 71 | PRIOR FILING DATE: 1998-06-26 |
| 72 | PRIOR APPLICATION NUMBER: 60/090863 |
| 73 | PRIOR FILING DATE: 1998-06-26 |

```
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      100.0%; Score 20; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
        1111
Db      21 AAPM 24

RESULT 35
US-09-993-604-401
; Sequence 401, Application US/09993604
; Patent No. US20020137075A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C25
; CURRENT APPLICATION NUMBER: US/09/993,604
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
```

```
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
```

;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519

;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09
Query Match 100.0%; Score 20; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 APM 4
1111
DB 21 APM 24
RESULT 36
US-09-990-456-401
; Sequence 401, Application US/09990456
; Patent No. US20020137890A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Bolstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C22
; CURRENT APPLICATION NUMBER: US/09/990,456
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07

| | | | | | |
|----|-------------------------------------|----|-------------------------------------|----|-------------------------------------|
| 1 | PRIOR APPLICATION NUMBER: 60/087106 | 1 | PRIOR FILING DATE: 1998-06-17 | 1 | PRIOR FILING DATE: 1998-07-07 |
| 2 | PRIOR FILING DATE: 1998-05-28 | 2 | PRIOR APPLICATION NUMBER: 60/089801 | 2 | PRIOR FILING DATE: 1998-07-02 |
| 3 | PRIOR APPLICATION NUMBER: 60/087607 | 3 | PRIOR FILING DATE: 1998-06-18 | 3 | PRIOR APPLICATION NUMBER: 60/091633 |
| 4 | PRIOR FILING DATE: 1998-06-02 | 4 | PRIOR APPLICATION NUMBER: 60/089907 | 4 | PRIOR FILING DATE: 1998-07-02 |
| 5 | PRIOR APPLICATION NUMBER: 60/087609 | 5 | PRIOR FILING DATE: 1998-06-18 | 5 | PRIOR APPLICATION NUMBER: 60/091978 |
| 6 | PRIOR FILING DATE: 1998-06-02 | 6 | PRIOR APPLICATION NUMBER: 60/089908 | 6 | PRIOR FILING DATE: 1998-07-07 |
| 7 | PRIOR APPLICATION NUMBER: 60/087759 | 7 | PRIOR FILING DATE: 1998-06-18 | 7 | PRIOR FILING DATE: 1998-07-02 |
| 8 | PRIOR FILING DATE: 1998-06-02 | 8 | PRIOR APPLICATION NUMBER: 60/089948 | 8 | PRIOR APPLICATION NUMBER: 60/089600 |
| 9 | PRIOR APPLICATION NUMBER: 60/087827 | 9 | PRIOR FILING DATE: 1998-06-19 | 9 | PRIOR FILING DATE: 1998-06-17 |
| 10 | PRIOR FILING DATE: 1998-06-03 | 10 | PRIOR APPLICATION NUMBER: 60/089952 | 10 | PRIOR APPLICATION NUMBER: 60/089598 |
| 11 | PRIOR APPLICATION NUMBER: 60/088021 | 11 | PRIOR FILING DATE: 1998-06-19 | 11 | PRIOR FILING DATE: 1998-06-17 |
| 12 | PRIOR FILING DATE: 1998-06-04 | 12 | PRIOR APPLICATION NUMBER: 60/090246 | 12 | PRIOR APPLICATION NUMBER: 60/089599 |
| 13 | PRIOR APPLICATION NUMBER: 60/088025 | 13 | PRIOR FILING DATE: 1998-06-19 | 13 | PRIOR FILING DATE: 1998-06-17 |
| 14 | PRIOR FILING DATE: 1998-06-04 | 14 | PRIOR APPLICATION NUMBER: 60/090252 | 14 | PRIOR APPLICATION NUMBER: 60/089653 |
| 15 | PRIOR APPLICATION NUMBER: 60/088026 | 15 | PRIOR FILING DATE: 1998-06-22 | 15 | PRIOR FILING DATE: 1998-06-17 |
| 16 | PRIOR FILING DATE: 1998-06-04 | 16 | PRIOR APPLICATION NUMBER: 60/090254 | 16 | PRIOR FILING DATE: 1998-06-17 |
| 17 | PRIOR APPLICATION NUMBER: 60/088028 | 17 | PRIOR FILING DATE: 1998-06-22 | 17 | PRIOR FILING DATE: 1998-06-17 |
| 18 | PRIOR FILING DATE: 1998-06-04 | 18 | PRIOR APPLICATION NUMBER: 60/090349 | 18 | PRIOR FILING DATE: 1998-06-17 |
| 19 | PRIOR APPLICATION NUMBER: 60/088029 | 19 | PRIOR FILING DATE: 1998-06-23 | 19 | PRIOR FILING DATE: 1998-06-17 |
| 20 | PRIOR FILING DATE: 1998-06-04 | 20 | PRIOR APPLICATION NUMBER: 60/090355 | 20 | PRIOR FILING DATE: 1998-06-17 |
| 21 | PRIOR APPLICATION NUMBER: 60/088030 | 21 | PRIOR FILING DATE: 1998-06-23 | 21 | PRIOR FILING DATE: 1998-06-17 |
| 22 | PRIOR FILING DATE: 1998-06-04 | 22 | PRIOR APPLICATION NUMBER: 60/090429 | 22 | PRIOR FILING DATE: 1998-06-17 |
| 23 | PRIOR APPLICATION NUMBER: 60/088033 | 23 | PRIOR FILING DATE: 1998-06-24 | 23 | PRIOR FILING DATE: 1998-06-17 |
| 24 | PRIOR FILING DATE: 1998-06-04 | 24 | PRIOR APPLICATION NUMBER: 60/090431 | 24 | PRIOR FILING DATE: 1998-06-17 |
| 25 | PRIOR APPLICATION NUMBER: 60/088326 | 25 | PRIOR FILING DATE: 1998-06-24 | 25 | PRIOR FILING DATE: 1998-06-17 |
| 26 | PRIOR FILING DATE: 1998-06-04 | 26 | PRIOR APPLICATION NUMBER: 60/090435 | 26 | PRIOR FILING DATE: 1998-06-17 |
| 27 | PRIOR APPLICATION NUMBER: 60/088167 | 27 | PRIOR FILING DATE: 1998-06-24 | 27 | PRIOR FILING DATE: 1998-06-17 |
| 28 | PRIOR FILING DATE: 1998-06-05 | 28 | PRIOR APPLICATION NUMBER: 60/090444 | 28 | PRIOR FILING DATE: 1998-06-17 |
| 29 | PRIOR APPLICATION NUMBER: 60/088202 | 29 | PRIOR FILING DATE: 1998-06-24 | 29 | PRIOR FILING DATE: 1998-06-17 |
| 30 | PRIOR FILING DATE: 1998-06-05 | 30 | PRIOR APPLICATION NUMBER: 60/090445 | 30 | PRIOR FILING DATE: 1998-06-17 |
| 31 | PRIOR APPLICATION NUMBER: 60/088212 | 31 | PRIOR FILING DATE: 1998-06-24 | 31 | PRIOR FILING DATE: 1998-06-17 |
| 32 | PRIOR FILING DATE: 1998-06-05 | 32 | PRIOR APPLICATION NUMBER: 60/090472 | 32 | PRIOR FILING DATE: 1998-06-17 |
| 33 | PRIOR APPLICATION NUMBER: 60/088217 | 33 | PRIOR FILING DATE: 1998-06-24 | 33 | PRIOR FILING DATE: 1998-06-17 |
| 34 | PRIOR FILING DATE: 1998-06-05 | 34 | PRIOR APPLICATION NUMBER: 60/090535 | 34 | PRIOR FILING DATE: 1998-06-17 |
| 35 | PRIOR APPLICATION NUMBER: 60/088655 | 35 | PRIOR FILING DATE: 1998-06-24 | 35 | PRIOR FILING DATE: 1998-06-17 |
| 36 | PRIOR FILING DATE: 1998-06-09 | 36 | PRIOR APPLICATION NUMBER: 60/090540 | 36 | PRIOR FILING DATE: 1998-06-17 |
| 37 | PRIOR APPLICATION NUMBER: 60/088734 | 37 | PRIOR FILING DATE: 1998-06-24 | 37 | PRIOR FILING DATE: 1998-06-17 |
| 38 | PRIOR FILING DATE: 1998-06-10 | 38 | PRIOR APPLICATION NUMBER: 60/090542 | 38 | PRIOR FILING DATE: 1998-06-17 |
| 39 | PRIOR APPLICATION NUMBER: 60/088738 | 39 | PRIOR FILING DATE: 1998-06-24 | 39 | PRIOR FILING DATE: 1998-06-17 |
| 40 | PRIOR FILING DATE: 1998-06-10 | 40 | PRIOR APPLICATION NUMBER: 60/090557 | 40 | PRIOR FILING DATE: 1998-06-17 |
| 41 | PRIOR APPLICATION NUMBER: 60/088742 | 41 | PRIOR FILING DATE: 1998-06-24 | 41 | PRIOR FILING DATE: 1998-06-17 |
| 42 | PRIOR FILING DATE: 1998-06-10 | 42 | PRIOR APPLICATION NUMBER: 60/090676 | 42 | PRIOR FILING DATE: 1998-06-17 |
| 43 | PRIOR APPLICATION NUMBER: 60/088810 | 43 | PRIOR FILING DATE: 1998-06-25 | 43 | PRIOR FILING DATE: 1998-06-17 |
| 44 | PRIOR FILING DATE: 1998-06-10 | 44 | PRIOR APPLICATION NUMBER: 60/090678 | 44 | PRIOR FILING DATE: 1998-06-17 |
| 45 | PRIOR APPLICATION NUMBER: 60/088824 | 45 | PRIOR FILING DATE: 1998-06-25 | 45 | PRIOR FILING DATE: 1998-06-17 |
| 46 | PRIOR FILING DATE: 1998-06-10 | 46 | PRIOR APPLICATION NUMBER: 60/090690 | 46 | PRIOR FILING DATE: 1998-06-17 |
| 47 | PRIOR APPLICATION NUMBER: 60/088826 | 47 | PRI | | |

;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 20; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
|||

Db 21 AAPM 24

RESULT 37

US-09-989-721-401
; Sequence 401, Application US/09989721
; Patent No. US20020142961A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C55
; CURRENT APPLICATION NUMBER: US/09/989,721
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759

;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18

| | | |
|----|--------------------|-------------------|
| 1 | PRIOR APPLICATION | NUMBER: 60/089947 |
| 2 | PRIOR FILING DATE: | 1998-06-19 |
| 3 | PRIOR APPLICATION | NUMBER: 60/089948 |
| 4 | PRIOR FILING DATE: | 1998-06-19 |
| 5 | PRIOR APPLICATION | NUMBER: 60/089952 |
| 6 | PRIOR FILING DATE: | 1998-06-19 |
| 7 | PRIOR APPLICATION | NUMBER: 60/090246 |
| 8 | PRIOR FILING DATE: | 1998-06-22 |
| 9 | PRIOR APPLICATION | NUMBER: 60/090252 |
| 10 | PRIOR FILING DATE: | 1998-06-22 |
| 11 | PRIOR APPLICATION | NUMBER: 60/090254 |
| 12 | PRIOR FILING DATE: | 1998-06-22 |
| 13 | PRIOR APPLICATION | NUMBER: 60/090349 |
| 14 | PRIOR FILING DATE: | 1998-06-23 |
| 15 | PRIOR APPLICATION | NUMBER: 60/090355 |
| 16 | PRIOR FILING DATE: | 1998-06-23 |
| 17 | PRIOR APPLICATION | NUMBER: 60/090429 |
| 18 | PRIOR FILING DATE: | 1998-06-24 |
| 19 | PRIOR APPLICATION | NUMBER: 60/090431 |
| 20 | PRIOR FILING DATE: | 1998-06-24 |
| 21 | PRIOR APPLICATION | NUMBER: 60/090435 |
| 22 | PRIOR FILING DATE: | 1998-06-24 |
| 23 | PRIOR APPLICATION | NUMBER: 60/090444 |
| 24 | PRIOR FILING DATE: | 1998-06-24 |
| 25 | PRIOR APPLICATION | NUMBER: 60/090445 |
| 26 | PRIOR FILING DATE: | 1998-06-24 |
| 27 | PRIOR APPLICATION | NUMBER: 60/090472 |
| 28 | PRIOR FILING DATE: | 1998-06-24 |
| 29 | PRIOR APPLICATION | NUMBER: 60/090535 |
| 30 | PRIOR FILING DATE: | 1998-06-24 |
| 31 | PRIOR APPLICATION | NUMBER: 60/090540 |
| 32 | PRIOR FILING DATE: | 1998-06-24 |
| 33 | PRIOR APPLICATION | NUMBER: 60/090542 |
| 34 | PRIOR FILING DATE: | 1998-06-24 |
| 35 | PRIOR APPLICATION | NUMBER: 60/090557 |
| 36 | PRIOR FILING DATE: | 1998-06-24 |
| 37 | PRIOR APPLICATION | NUMBER: 60/090676 |
| 38 | PRIOR FILING DATE: | 1998-06-25 |
| 39 | PRIOR APPLICATION | NUMBER: 60/090678 |
| 40 | PRIOR FILING DATE: | 1998-06-25 |
| 41 | PRIOR APPLICATION | NUMBER: 60/090690 |
| 42 | PRIOR FILING DATE: | 1998-06-25 |
| 43 | PRIOR APPLICATION | NUMBER: 60/090694 |
| 44 | PRIOR FILING DATE: | 1998-06-25 |
| 45 | PRIOR APPLICATION | NUMBER: 60/090695 |
| 46 | PRIOR FILING DATE: | 1998-06-25 |
| 47 | PRIOR APPLICATION | NUMBER: 60/090696 |
| 48 | PRIOR FILING DATE: | 1998-06-25 |
| 49 | PRIOR APPLICATION | NUMBER: 60/090862 |
| 50 | PRIOR FILING DATE: | 1998-06-26 |
| 51 | PRIOR APPLICATION | NUMBER: 60/090863 |
| 52 | PRIOR FILING DATE: | 1998-06-26 |
| 53 | PRIOR APPLICATION | NUMBER: 60/091360 |
| 54 | PRIOR FILING DATE: | 1998-07-01 |
| 55 | PRIOR APPLICATION | NUMBER: 60/091478 |
| 56 | PRIOR FILING DATE: | 1998-07-02 |
| 57 | PRIOR APPLICATION | NUMBER: 60/091544 |
| 58 | PRIOR FILING DATE: | 1998-07-01 |
| 59 | PRIOR APPLICATION | NUMBER: 60/091519 |
| 60 | PRIOR FILING DATE: | 1998-07-02 |
| 61 | PRIOR APPLICATION | NUMBER: 60/091626 |
| 62 | PRIOR FILING DATE: | 1998-07-02 |
| 63 | PRIOR APPLICATION | NUMBER: 60/091633 |
| 64 | PRIOR FILING DATE: | 1998-07-02 |
| 65 | PRIOR APPLICATION | NUMBER: 60/091978 |
| 66 | PRIOR FILING DATE: | 1998-07-07 |
| 67 | PRIOR APPLICATION | NUMBER: 60/091982 |
| 68 | PRIOR FILING DATE: | 1998-07-07 |
| 69 | PRIOR APPLICATION | NUMBER: 60/092182 |
| 70 | PRIOR FILING DATE: | 1998-07-09 |

| Query Match | Score 20; | DB 10; | Length 198; |
|-----------------------|-----------|--------------------|-------------|
| Best Local Similarity | 100.0%; | Pred. No. 2.5e+02; | |

| | | | | | | | | | |
|---------|----|--------------|----|------------|----|--------|----|------|----|
| Matches | 4; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
| QY | 1 | AAPM | 4 | | | | | | |
| | | | | | | | | | |
| Db | 21 | AAPM | 24 | | | | | | |

```

RESULT 38
; US-09-244-694-3
; Sequence 3, Application US/09244694
; Patent No. US20020026037A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Jing-Shan
; APPLICANT: Olsen, Henrik S.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: Vascular Endothelial Growth Factor 3
; FILE REFERENCE: 1488.1040003
; CURRENT APPLICATION NUMBER: US/09/244,694
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: US 09/132,088
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: US 09/033,662
; EARLIER FILING DATE: 1998-03-03
; EARLIER APPLICATION NUMBER: US 08/469,641
; EARLIER FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-244-694-3

```

| | | | | |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match | 100.0%; | Score 20; | DB 10; | Length 215; |
| Best Local Similarity | 100.0%; | Pred. No. 2.7e+02; | | |
| Matches 4; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

| | | | |
|----|----|------|----|
| QY | 1 | AAPM | 4 |
| | | | |
| Db | 26 | AAPM | 29 |

```

RESULT 39
US-09-291-299A-8
; Sequence 8, Application US/09291299A
; Patent No. US20020076405A1
; GENERAL INFORMATION:
; APPLICANT: Yong, Xie
; TITLE OF INVENTION: Antigenized Antibody Vaccine for Foot-and-Mouth Disease
; FILE REFERENCE: 09/291,299
; CURRENT APPLICATION NUMBER: US/09/291,299A
; CURRENT FILING DATE: 1999-04-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Sus sp.
US-09-291-299A-8

```

| | | | | |
|-----------------------|--------------|--------------------|--------|---------------|
| Query Match | 100.0%; | Score 20; | DB 10; | Length 216; |
| Best Local Similarity | 100.0%; | Pred. No. 2.7e+02; | | |
| Matches | 4; | Conservative | 0; | Mismatches 0; |
| | | | | Indels 0; |
| | | | | Gaps 0; |
| QY | 1 AAPM 4 | | | |
| | | | | |
| DB | 115 AAPM 118 | | | |

RESULT 40
US-09-925-299-967
; Sequence 967, Application US/09925299
; Patent No. US20020055627A1

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 967
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-967

Query Match 100.0%; Score 20; DB 10; Length 221;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 48 AAPM 51

Search completed: December 6, 2002, 13:42:06
Job time : 5.88889 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:25:59 ; Search time 10.2222 Seconds
(without alignments)
37.618 Million cell updates/sec

Title: US-10-033-526-2
Perfect score: 20
Sequence: 1 AAPM 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 20 | 100.0 | 41 | 1 B40236 | structural protein |
| 2 | 20 | 100.0 | 60 | 1 H3NJ1M | cytotoxin 1 - Moza |
| 3 | 20 | 100.0 | 60 | 1 H3NJ1B | cytotoxin 1 - cobr |
| 4 | 20 | 100.0 | 79 | 2 T30118 | hypothetical prote |
| 5 | 20 | 100.0 | 83 | 2 B95974 | probable transcrip |
| 6 | 20 | 100.0 | 88 | 2 F97400 | hypothetical prote |
| 7 | 20 | 100.0 | 102 | 1 T1SYD2 | proteinase inhibit |
| 8 | 20 | 100.0 | 105 | 2 T36139 | hypothetical prote |
| 9 | 20 | 100.0 | 105 | 2 C72580 | hypothetical prote |
| 10 | 20 | 100.0 | 106 | 2 G97841 | acetyl-CoA acetyl |
| 11 | 20 | 100.0 | 108 | 2 C97545 | conserved hypotet |
| 12 | 20 | 100.0 | 108 | 2 AD2764 | hypothetical prote |
| 13 | 20 | 100.0 | 109 | 2 F82556 | hypothetical prote |
| 14 | 20 | 100.0 | 121 | 2 T22303 | hypothetical prote |
| 15 | 20 | 100.0 | 121 | 2 D75584 | hypothetical prote |
| 16 | 20 | 100.0 | 123 | 1 VHWME | structural protein |
| 17 | 20 | 100.0 | 123 | 1 C44212 | structural protein |
| 18 | 20 | 100.0 | 126 | 2 C86883 | 50S ribosomal prot |
| 19 | 20 | 100.0 | 128 | 2 H95027 | 50S ribosomal prot |
| 20 | 20 | 100.0 | 128 | 2 H97898 | isoprenylated prot |
| 21 | 20 | 100.0 | 130 | 2 T06394 | conserved hypotet |
| 22 | 20 | 100.0 | 131 | 2 AC0747 | probable holo-acyl |
| 23 | 20 | 100.0 | 133 | 2 F75544 | conserved hypotet |
| 24 | 20 | 100.0 | 139 | 2 D87701 | conserved hypotet |
| 25 | 20 | 100.0 | 140 | 2 T01170 | ferredoxin [2Fe-2S |
| 26 | 20 | 100.0 | 142 | 2 G64952 | hypothetical 17.1 |
| 27 | 20 | 100.0 | 142 | 2 C90954 | probable regulator |
| 28 | 20 | 100.0 | 142 | 2 H85802 | probable regulator |
| 29 | 20 | 100.0 | 146 | 2 S57956 | ovine vascular end |

| | | | | | |
|----|----|-------|-----|----------|---------------------|
| 30 | 20 | 100.0 | 147 | 2 AG3578 | hypothetical prote |
| 31 | 20 | 100.0 | 152 | 2 G98270 | tolr protein (tolr |
| 32 | 20 | 100.0 | 152 | 2 AH3013 | tolr protein [limp |
| 33 | 20 | 100.0 | 152 | 2 C75544 | hypothetical prote |
| 34 | 20 | 100.0 | 154 | 2 T06396 | isoprenylated prot |
| 35 | 20 | 100.0 | 156 | 1 BREC9 | acetyl-CoA carboxy |
| 36 | 20 | 100.0 | 156 | 2 T49921 | ribosomal protein-1 |
| 37 | 20 | 100.0 | 156 | 2 C85990 | acetyl-CoA carboxy |
| 38 | 20 | 100.0 | 156 | 2 G91144 | acetyl-CoA carboxy |
| 39 | 20 | 100.0 | 156 | 2 AH0912 | blotin carboxyl ca |
| 40 | 20 | 100.0 | 159 | 2 D87658 | hypothetical prote |
| 41 | 20 | 100.0 | 159 | 2 A82219 | transcription regu |
| 42 | 20 | 100.0 | 159 | 2 T35614 | polyketide cyclase |
| 43 | 20 | 100.0 | 170 | 2 B72122 | ribosomal protein |
| 44 | 20 | 100.0 | 170 | 2 G86500 | L10 ribosomal prot |
| 45 | 20 | 100.0 | 171 | 2 S35248 | nifq protein - Ent |

ALIGNMENTS

RESULT 1
B40236
structural protein 1 - hepatitis E virus (strain Indian)
C:Species: hepatitis E virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C:Accession: B40236
R:Ray, R.; Jameel, S.; Manivel, V.; Ray, R.
Virology 189, 359-362, 1992
A:Title: Indian hepatitis E virus shows a major deletion in the small open reading fr
A:Reference number: A40236; MUID:92295577; PMID:1534953
A:Accession: B40236
A:Molecule type: genomic RNA
A:Residues: 1-41 <RAY>
C:Superfamily: hepatitis E virus structural protein 1
C:Keywords: structural protein

Query Match 100.0%; Score 20; DB 1; Length 41;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
DB 7 AAPM 10

RESULT 2
H3NJ1M
cytotoxin 1 - Mozambique cobra (tentative sequence)
N:Alternate names: cardiotoxin gamma; cardiotoxin XIIB; cytotoxin XIIB
C:Species: Naja mossambica mossambica (Mozambique cobra)
C:Date: 29-Jun-1981 #sequence_revision 29-Jun-1981 #text_change 31-Mar-2000
C:Accession: A38050; A01728; S02517
R:Low, A.I.
Biochim. Biophys. Acta 336, 481-495, 1974
A:Title: Snake venom toxins. The amino acid sequences of three cytotoxin homologues f
A:Reference number: A01728
A:Accession: A38050
A:Molecule type: protein
A:Residues: 1-60 <LOU>
R:Otting, G.; Steinmetz, W.E.; Bougis, P.E.; Rochat, H.; Wuethrich, K.
Eur. J. Biochem. 168, 609-620, 1987
A:Title: Sequence-specific (1)H-NMR assignments and determination of the secondary st
A:Reference number: S02517; MUID:88029481; PMID:2822421
A:Contents: annotation; solution structure by NMR
C:Superfamily: snake toxin
C:Keywords: cytotoxin
F;3-21,14-38,42-53,54-59/Disulfide bonds: #status predicted
Query Match 100.0%; Score 20; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
|||
Db 28 AAPM 31

RESULT 3
H3NJ1B

cytotoxin 1 - cobra (Naja mossambica pallida) (tentative sequence)
N:Alternate names: cardiotoxin gamma
C:Species: Naja mossambica pallida
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 31-Mar-2000
C:Accession: A37578; A01728
R:Rycklund, L.; Eaker, D.
Biochemistry 14, 2865-2871, 1975
A:Title: The complete covalent structure of a cardiotoxin from the venom of Naja nigricollis
A:Reference number: A37578; MUID:75205552; PMID:1148181
A:Accession: A37578
A:Molecule type: protein
A:Residues: 1-60 <FRY>
C:Superfamily: snake toxin
C:Keywords: cytotoxin; hemolysis
F:3-21,14-38,42-53,54-59/Disulfide bonds: #status experimental

Query Match 100.0%; Score 20; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
|||
Db 28 AAPM 31

RESULT 4
T30118

hypothetical protein F22H10.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T30118
R:Langston, Y.; Hawkins, J.
submitted to the EMBL Data Library, September 1996
A:Description: The sequence of C. elegans cosmid F22H10.
A:Reference number: Z20740
A:Accession: T30118
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-79 <LAN>
A:Cross-references: EMBL:U70845; PIDN:AAB09098.1; GSPDB:GN00028; CESP:F22H10.3
A:Experimental source: strain Bristol N2; clone F22H10
C:Genetics:
A:Gene: CESP:F22H10.3
A:Map position: X

Query Match 100.0%; Score 20; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
|||
Db 47 AAPM 50

RESULT 5
B95974

probable transcription regulator protein [imported] - Sinorhizobium meliloti (strain 102)
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: B95974
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: B95974
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-83 <KUR>
A:Cross-references: GB:AL591985; PIDN:CAC49458.1; PID:g15140944; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubli
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation

C:Genetics:
A:Gene: exsI; SMB20935
A:Genome: plasmid

Query Match 100.0%; Score 20; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
|||
Db 13 AAPM 16

RESULT 6
F97400

hypothetical protein AGR_C_597 [imported] - Agrobacterium tumefaciens (strain C58, Ce
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: F97400
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.: Liu, F.; Wollam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: F97400
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-88 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK86159.1; PID:g15155250; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_597
A:Map position: circular chromosome

Query Match 100.0%; Score 20; DB 2; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
|||
Db 64 AAPM 67

RESULT 7
T1SYD2

proteinase inhibitor (Bowman-Birk) D-II precursor - soybean
N:Contains: proteinase inhibitor (Bowman-Birk) E-I (PI-II)
C:Species: Glycine max (soybean)
C:Date: 24-Apr-1984 #sequence_revision 12-Apr-1996 #text_change 18-Jun-1999
C:Accession: S32243; J02224; A01303; B01303; S29559
R:Song, J.C.; Baek, J.M.; Kim, S.I.
submitted to the EMBL Data Library, October 1992
A:Description: Molecular cloning of a gemonic DNA encoding the soybean bowman-birk pr
A:Reference number: S32243
A:Accession: S32243
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-102 <SON>
A:Cross-references: EMBL:X68707; NID:g288618; PIDN:CAA48658.1; PID:g288619
R:Baek, J.M.; Song, J.C.; Choi, Y.D.; Kim, S.I.
Biosci. Biotechnol. Biochem. 58, 843-846, 1994

A:Title: Nucleotide sequence homology of cDNAs encoding soybean Bowman-Birk type protein
A:Reference number: JC2224; MUID:94289861; PMID:7764974
A:Accession: JC2224
A:Molecule type: mRNA
A:Residues: 1-102 <BA2>
A:Cross-references: EMBL:X68706; NID:g18571; PIDN:CAA48657.1; PID:g18572
A:Experimental source: clone pB26
R:Odani, S.; Ikenaka, T.
J. Biochem. 83, 737-745, 1978
A:Title: Studies on soybean trypsin inhibitors. XII. Linear sequences of two soybean dou
A:Reference number: A01303; MUID:78150870; PMID:641033
A:Accession: A01303
A:Molecule type: protein
A:Residues: 28-102 <OD1>
A:Accession: B01303
A:Molecule type: protein
A:Residues: 37-57 <OD2>
R:Chen, P.; Rose, J.; Love, R.; Wei, C.H.; Wang, B.C.
J. Biol. Chem. 267, 1990-1994, 1992
A:Title: Reactive sites of an anticarcinogenic Bowman-Birk proteinase inhibitor are simi
A:Reference number: A42052; MUID:92112932; PMID:1730730
A:Contents: annotation; X-ray crystallography of inhibitor PI-II at 2.5 angstroms
C:Comment: This protein regulates endogenous proteinase during germination, stores sulf
C:Superfamily: Bowman-Birk proteinase inhibitor; Bowman-Birk inhibitor repeat homology
C:Keywords: duplication; seed; serine proteinase inhibitor; storage protein
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-102/Product: proteinase inhibitor (Bowman-Birk) D-II #status experimental <MAT>
F:44-70/Domain: Bowman-Birk inhibitor repeat homology <BB1>
F:71-96/Domain: Bowman-Birk inhibitor repeat homology <BB2>
F:43-97,44-59,47-93,49-57,67-74,71-86,76-84/Disulfide bonds: #status experimental
F:51/inhibitory site: Arg (trypsin) #status predicted
F:78/inhibitory site: Arg (trypsin) #status predicted

Query Match 100.0%; Score 20; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
 |||||
DB 18 AAPM 21

RESULT 8
T36139
hypothetical protein SCE19A.16c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36139
R:Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z21598
A:Accession: T36139
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-105 <SEE>
A:Cross-references: EMBL:AL096852; PIDN:CAB50997.1; GSPDB:GN00070; SCOEDB:SCE19A.16c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCE19A.16c

Query Match 100.0%; Score 20; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
 |||||
DB 7 AAPM 10

RESULT 9
C72580
hypothetical protein APE1923 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: C72580
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: C72580
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-105 <KAW>
A:Cross-references: DDBJ:AP000062; NID:g5105244; PIDN:BAA80928.1; PID:g1044714; PID:g
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1923

Query Match 100.0%; Score 20; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
 |||||
DB 43 AAPM 46

RESULT 10
G97841
acetyl-CoA acetyltransferase homolog RC1135 [imported] - Rickettsia conorii (strain M
C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C:Accession: G97841
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: G97841
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-106 <KUR>
A:Cross-references: GB:AE006914; PIDN:AAL03673.1; PID:g15620261; GSPDB:GN00173
C:Genetics:
A:Gene: RC1135

Query Match 100.0%; Score 20; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
 |||||
DB 26 AAPM 29

RESULT 11
C97545
hypothetical protein AGR_C_2812 [imported] - Agrobacterium tumefaciens (strain C58, C
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: C97545
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: C97545
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87316.1; PID:g15156613; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_2812
A:Map position: circular chromosome

Query Match 100.0%; Score 20; DB 2; Length 108;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 15 AAPM 18

RESULT 12

AD2764

conserved hypothetical protein Atu1525 [imported] - Agrobacterium tumefaciens (strain C58)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: AD2764

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AD2764

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-108 <KUR>

A:Cross-references: GB:AE008688; PIDN:AAL42530.1; PID:g17739951; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu1525

A:Map position: circular chromosome

Query Match 100.0%; Score 20; DB 2; Length 108;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 15 AAPM 18

RESULT 13

F82556

hypothetical protein XF2441 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: F82556

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: F82556

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-109 <SIM>

A:Cross-references: GB:AE004053; GB:AE003849; NID:g9107631; PIDN:AAF85240.1; GSPDB:GN001

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Rehnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, F

; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF2441

Query Match

100.0%; Score 20; DB 2; Length 109;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 36 AAPM 39

RESULT 14

T22303

hypothetical protein F46C5.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T22303

R:Thomas, K.

submitted to the EMBL Data Library, September 1995

A:Reference number: Z19544

A:Accession: T22303

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-121 <WIL>

A:Cross-references: EMBL:Z54281; PIDN:CAA91048.1; GSPDB:GN00020; CESP:F46C5.1

A:Experimental source: clone F46C5

C:Genetics:

A:Gene: CESP:F46C5.1

A:Map position: 2

A:Introns: 17/1; 109/3

C:Superfamily: Caenorhabditis elegans hypothetical protein F46C5.1

Query Match 100.0%; Score 20; DB 2; Length 121;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 22 AAPM 25

RESULT 15

D75584

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000

C:Accession: D75584

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: D75584

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-121 <WHI>

A:Cross-references: GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF12515.1; PID:g646

A:Experimental source: strain R1

C:Genetics:

A:Gene: DRA0303

A:Map position: 2

C:Superfamily: Deinococcus radiodurans hypothetical protein DRA0303

Query Match 100.0%; Score 20; DB 2; Length 121;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 40 AAPM 43

```
RESULT 16
VHWMHE
structural protein 1 - hepatitis E virus (strain Burma)
C/Species: hepatitis E virus
C/Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 23-Jul-1999
C/Accession: B40778; A40236
R/Tam, A.W.; Smith, M.M.; Guerra, M.E.; Huang, C.C.; Bradley, D.W.; Fry, K.E.; Reyes, G.
Virology 185, 120-131, 1991
A/Title: Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length vi
A/Reference number: A40778; MUID:92024067; PMID:1926770
A/Accession: B40778
A/Molecule type: genomic RNA
A/Residues: 1-123 <TAM>
A/Cross-references: GB:M73218; NID:g330023; PIDN:AAA45735.1; PID:g330025
R/Ray, R.; Jameel, S.; Manivel, V.; Ray, R.
Virology 189, 359-362, 1992
A/Title: Indian hepatitis E virus shows a major deletion in the small open reading frame
A/Reference number: A40236; MUID:92295577; PMID:1534953
A/Accession: A40236
A/Molecule type: genomic RNA
A/Residues: 1-57, 'P', 59-102, 'P', 104-123 <RAY>
C/Superfamily: hepatitis E virus structural protein 1
C/Keywords: structural protein

Query Match          100.0%; Score 20; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 AAPM 4
        ||||
Db       7 AAPM 10

RESULT 17
C44212
structural protein 1 - hepatitis E virus (strain Mexico)
C/Species: hepatitis E virus
C/Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C/Accession: C44212
R/Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.
Virology 191, 550-558, 1992
A/Title: Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HE
A/Reference number: A44212; MUID:93079857; PMID:1448913
A/Accession: C44212
A/Molecule type: genomic RNA
A/Residues: 1-123 <HUA>
A/Cross-references: GB:M74506; NID:g330017; PIDN:AAA45731.1; PID:g330019
C/Superfamily: hepatitis E virus structural protein 1
C/Keywords: structural protein

Query Match          100.0%; Score 20; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 AAPM 4
        ||||
Db       7 AAPM 10

RESULT 18
C86883
50S ribosomal protein L17 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C/Accession: C86883
R/Bolotin, A.; Wincker, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: C86883
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-126 <STO>
```

```
A/Cross-references: GB:AE005176; PID:g12725119; PIDN:AAK06165.1; GSPDB:GN00146
A/Experimental source: strain IL1403
C/Genetics:
A/Gene: rplQ
C/Superfamily: Escherichia coli ribosomal protein L17

Query Match          100.0%; Score 20; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 AAPM 4
        ||||
Db       117 AAPM 120

RESULT 19
H95027
ribosomal protein L17 [imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C/Accession: H95027
R/Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtappl
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MUID:21357209; PMID:11463916
A/Accession: H95027
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-128 <KUR>
A/Cross-references: GB:AE005672; PIDN:AAK74417.1; PID:g14971706; GSPDB:GN00164; TIGR:
A/Experimental source: strain TIGR4
C/Genetics:
A/Gene: SP0237
C/Superfamily: Escherichia coli ribosomal protein L17

Query Match          100.0%; Score 20; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 AAPM 4
        ||||
Db       119 AAPM 122

RESULT 20
H97898
50S ribosomal protein L17 [imported] - Streptococcus pneumoniae (strain R6)
C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C/Accession: H97898
R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A/Reference number: A97872; MUID:21429245; PMID:11544234
A/Accession: H97898
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-128 <KUR>
A/Cross-references: GB:AE007317; PIDN:AAK99020.1; PID:g15457761; GSPDB:GN00174
C/Genetics:
A/Gene: rplQ
C/Superfamily: Escherichia coli ribosomal protein L17

Query Match          100.0%; Score 20; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 AAPM 4
||||
Db 119 AAPM 122

RESULT 21

T06394

isoprenylated protein - soybean (fragment)

C:Species: Glycine max (soybean)

C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-May-2000

C:Accession: T06394

R:Biermann, B.J.; Morehead, T.A.; Tate, S.E.; Price, J.R.; Randall, S.K.; Crowell, D.N.
J. Biol. Chem. 269, 25251-25254, 1994

A:Title: Novel isoprenylated proteins identified by an expression library screen.

A:Reference number: Z15647; MUID:95014311; PMID:7929216

A:Accession: T06394

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-130 <BIE>

A:Cross-references: EMBL:U13180; NID:g532702; PIDN:AAA65012.1; PID:g532703

A:Experimental source: strain Mandarin

Query Match 100.0%; Score 20; DB 2; length 130;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||

Db 57 AAPM 60

RESULT 22

AC0747

conserved hypothetical protein STY2136 [imported] - Salmonella enterica subsp. enterica

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001

C:Accession: AC0747

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; PMID:11677608

A:Accession: AC0747

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-131 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD05678.1; PID:g16503173; GSPDB:GN00176

C:Genetics:

A:Gene: STY2136

C:Superfamily: universal stress protein A

Query Match 100.0%; Score 20; DB 2; length 131;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||

Db 50 AAPM 53

RESULT 23

F75544

probable holo-acyl carrier protein synthase - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: F75544

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: F75544

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-133 <WHI>

A:Cross-references: GB:AE00186; GB:AE000513; NID:g6457921; PIDN:AAF09839.1; PID:g645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0247

A:Map position: 1

C:Superfamily: holo-ACP synthase

Query Match 100.0%; Score 20; DB 2; length 133;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||

Db 86 AAPM 89

RESULT 24

D87701

conserved hypothetical protein CC3646 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: D87701

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: D87701

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-139 <STO>

A:Cross-references: GB:AE005673; NID:g13425402; PIDN:AAK25608.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC3646

Query Match 100.0%; Score 20; DB 2; length 139;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||

Db 71 AAPM 74

RESULT 25

T01170

ferredoxin [2Fe-2S] 2 - maize

C:Species: Zea mays (maize)

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 20-Jun-2000

C:Accession: T01170

R:Matsumura, T.; Kimata-Ariaga, Y.; Sakakibara, H.; Sugiyama, T.; Murata, H.; Takao, T

submitted to the EMBL Data Library, August 1998

A:Description: CDNA cloning and characterization of ferredoxin localized in bundle sh

A:Reference number: Z14252

A:Accession: T01170

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-140 <MAT>

C:Genetics:

A:Gene: PFD2

C:Superfamily: ferredoxin [2Fe-2S]; ferredoxin [2Fe-2S] homology

C:Keywords: 2Fe-2S; electron transfer; iron-sulfur protein; metalloprotein

F;68-122/Domain: ferredoxin [2Fe-2S] homology <FDX>

F;83,88,91,121/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

```
Query Match          100.0%; Score 20; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
    ||||
Db 33 AAPM 36

RESULT 26
G64952
hypothetical 17.1 kD protein in fhbD-otsA intergenic region - Escherichia coli (strain K
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: G64952
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: G64952
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-142 <BLAT>
A:Cross-references: GB:AE000283; GB:U00096; NID:g1788200; PIDN:AAC74965.1; PID:g1788205;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: yecG
C:Superfamily: universal stress protein A

Query Match          100.0%; Score 20; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
    ||||
Db 50 AAPM 53

RESULT 27
C90954
probable regulator [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: C90954
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C90954
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-142 <HAY>
A:Cross-references: GB:BA000007; PIDN:BAB36026.1; PID:g13362071; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs2603
C:Superfamily: universal stress protein A

Query Match          100.0%; Score 20; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
    ||||
Db 50 AAPM 53

RESULT 28
H85802
probable regulator yecG [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
```

```
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: H85802
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Llm, A.; Dimailanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85802
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-142 <STO>
A:Cross-references: GB:AE005174; NID:g12515958; PIDN:AAG56884.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yecG
C:Superfamily: universal stress protein A

Query Match          100.0%; Score 20; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
    ||||
Db 50 AAPM 53

RESULT 29
S57956
ovine vascular endothelial growth factor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-Nov-1999
C:Accession: S57956
R:Redmer, D.A.; Dai, Y.; Li, J.; Jones, S.C.; Moor, R.M.
submitted to the EMBL Data Library, July 1995
A:Reference number: S57956
A:Accession: S57956
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-146 <RED>
A:Cross-references: EMBL:X89506; NID:g899350; PIDN:CAA61677.1; PID:g899351

Query Match          100.0%; Score 20; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
    ||||
Db 26 AAPM 29

RESULT 30
AG3578
hypothetical protein BMEI10552 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AG3578
R:DelVecchio, V.G.; Kaparat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A:Reference number: AD3252; PMID:11756688
A:Accession: AG3578
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-147 <KUR>
A:Cross-references: GB:AE008918; PIDN:AAL53794.1; PID:g17984725; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI10552
A:Map position: II

Query Match          100.0%; Score 20; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
```


Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||
Db 26 AAPM 29

RESULT 31

G98270
tolR protein (tolR) rp310 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: G98270
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; PMID:11743194
A:Accession: G98270
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-152 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK89689.1; PID:g15159595; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_2238
A:Map position: linear chromosome
C:Superfamily: tolR protein

Query Match 100.0%; Score 20; DB 2; Length 152;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||
Db 49 AAPM 52

RESULT 32

AH3013
tolR protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: AH3013
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AH3013
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-152 <KUR>
A:Cross-references: GB:AE008689; PIDN:AAL44526.1; PID:g17742138; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: tolR
A:Map position: linear chromosome
C:Superfamily: tolR protein

Query Match 100.0%; Score 20; DB 2; Length 152;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||
Db 49 AAPM 52

RESULT 33

C75544

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: C75544
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Yamathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M. Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: C75544
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-152 <WHI>
A:Cross-references: GB:AE001886; GB:AE000513; NID:g6457921; PIDN:AAF09834.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0244
A:Map position: 1

Query Match 100.0%; Score 20; DB 2; Length 152;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||
Db 127 AAPM 130

RESULT 34

T06396
isoprenylated protein - soybean (fragment)
C:Species: Glycine max (soybean)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-May-2000
C:Accession: T06396
R:Biermann, B.J.; Morehead, T.A.; Tate, S.E.; Price, J.R.; Randall, S.K.; Crowell, D. J. Biol. Chem. 269, 25251-25254, 1994
A:Title: Novel isoprenylated proteins identified by an expression library screen.
A:Reference number: Z15647; MUID:95014311; PMID:7929216
A:Accession: T06396
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-154 <BIE>
A:Cross-references: EMBL:U13181; NID:g532704; PIDN:AAA65013.1; PID:g532705
A:Experimental source: strain Mandarin

Query Match 100.0%; Score 20; DB 2; Length 154;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||
Db 81 AAPM 84

RESULT 35

BKEC9
acetyl-CoA carboxylase (EC 6.4.1.2) biotin carboxyl carrier protein [validated] - Esc N:Contains: biotin carboxyl carrier protein
C:Species: Escherichia coli
C:Date: 30-Nov-1980 #sequence_revision 31-Dec-1989 #text_change 01-Mar-2002
C:Accession: A93687; A92204; JS0686; PS0357; A33643; C40637; PS0272; A65118; S78777; R:Muramatsu, S.; Mizuno, T. Nucleic Acids Res. 17, 3982, 1989
A:Title: Nucleotide sequence of the fabE gene and flanking regions containing a bent A:Reference number: A93687; MUID:89282408; PMID:2660106
A:Accession: A93687
A:Molecule type: DNA
A:Residues: 1-156 <MUR>
A:Cross-references: GB:X14825; NID:g41361; PIDN:CAA32933.1; PID:g41362
A:Experimental source: strain LA2-22
R:Sutton, M.R.; Fall, R.R.; Nevel, A.M.; Alberts, A.W.; Vagelos, P.R.; Bradshaw, R.A. J. Biol. Chem. 252, 3934-3940, 1977

A;Title: Amino acid sequence of Escherichia coli biotin carboxyl carrier protein (9100).
A;Reference number: A92204; MUID:77187896; PMID:324999
A;Accession: A92204
A;Molecule type: protein
A;Residues: 75-156 <SUT>
R;Li, S.J.; Cronan Jr., J.E.
J. Biol. Chem. 267, 855-863, 1992
A;Title: The gene encoding the biotin carboxylase subunit of Escherichia coli acetyl-CoA
A;Reference number: JS0686; MUID:92112819; PMID:1370469
A;Accession: JS0686
A;Molecule type: DNA
A;Residues: 1-156 <LIS>
A;Cross-references: GB:M80458; NID:g145172; PIDN:AAA23408.1; PID:g145174
A;Experimental source: strain K12
A;Accession: PS0357
A;Molecule type: protein
A;Residues: 1-23 <LISI>
R;Alix, J.H.
DNA 8, 779-789, 1989
A;Title: A rapid procedure for cloning genes from lambda libraries by complementation of
A;Reference number: A33643; MUID:90126231; PMID:2575489
A;Accession: A33643
A;Molecule type: DNA
A;Residues: 1-156 <ALI>
A;Cross-references: GB:M32214; NID:g145889; PIDN:AAA23744.1; PID:g145890
R;Li, S.J.; Cronan Jr., J.E.
J. Bacteriol. 175, 332-340, 1993
A;Title: Growth rate regulation of Escherichia coli acetyl coenzyme A carboxylase, which
A;Reference number: A40637; MUID:93123150; PMID:7678242
A;Accession: C40637
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-25 <LII>
A;Cross-references: GB:S52932; NID:g263399; PIDN:AAB24892.1; PID:g263401
A;Note: sequence extracted from NCBI backbone (NCBIN:122315, NCBI:P:122319)
R;Kondo, H.; Shiratsuchi, K.; Yoshimoto, T.; Masuda, T.; Kitazono, A.; Tsuru, D.; Anal,
Proc. Natl. Acad. Sci. U.S.A. 88, 9730-9733, 1991
A;Title: Acetyl-CoA carboxylase from Escherichia coli: gene organization and nucleotide
A;Reference number: JS0632; MUID:92052166; PMID:1682920
A;Accession: PS0272
A;Molecule type: DNA
A;Residues: 133-156 <KON>
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: A65118
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-156 <BIAT>
A;Cross-references: GB:AE000404; GB:U00096; NID:g2367207; PIDN:AAC76287.1; PID:g1789653;
A;Experimental source: strain K-12, substrain MGL655
R;Chapman-Smith, A.; Turner, D.L.; Cronan Jr., J.E.; Morris, T.W.; Wallace, J.C.
Biochem. J. 302, 881-887, 1994
A;Title: Expression, biotinylation and purification of a biotin-domain peptide from the
A;Reference number: S78777; MUID:95031932; PMID:7945216
A;Accession: S78777
A;Molecule type: protein
A;Residues: 2-20 <CHA>
C;Genetics:
A;Gene: accB; fabE
A;Map position: 72 min
C;Complex: in E. coli, acetyl-CoA carboxylase is composed of biotin carboxylase (EC 6.3.
biotin carboxyl carrier protein (BCCP, homodimer) (PIR:BKEC9)
C;Function: <ACO>
A;Description: EC 6.4.1.2 [validated; MUID:75035569]; the acetyl-CoA carboxylase complex
C;Function: <BCC>
A;Description: a specific lysine residue within the biotin carboxyl carrier protein (BCC
A;Note: the C-terminal 87 amino acids of the biotin carboxyl carrier protein (BCCP87) fo
C;Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology

C;Keywords: biotin binding; fatty acid biosynthesis; homodimer; ligase
F:1-156/Product: biotin carboxyl carrier protein BCCP, long form #status experimental
F:2-156/Product: biotin carboxyl carrier protein #status experimental <MAT>
F:70-156/Product: biotin carboxyl carrier protein BCCP87, short form #status experime
F:76-156/Domain: lipoyl/biotin-binding homology <LPB>
F:122/Binding site: biotin (Lys) (covalent) #status experimental

Query Match 100.0%; Score 20; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
DB 49 AAPM 52

RESULT 36
T49921

ribosomal protein-like - Arabidopsis thaliana
N:Alternate names: protein F17I14.40

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Sep-2000

C:Accession: T49921

R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancro
submitted to the Protein Sequence Database, April 2000

A;Reference number: Z24490

A;Accession: T49921

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-156 <BEV>

A;Cross-references: EMBL:AL353994; GSPDB:GN00063; ATSP:F17I14.40

A;Experimental source: cultivar Columbia; BAC clone F17I14

C;Genetics:
A;Gene: ATSP:F17I14.40

A;Map position: 5

A;Introns: 19/2; 37/3; 54/3; 85/2; 111/2

C;Superfamily: Escherichia coli ribosomal protein L17

Query Match 100.0%; Score 20; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
DB 104 AAPM 107

RESULT 37
C85990
acetyl-CoA carboxylase (EC 6.4.1.2) biotin carboxyl carrier protein [similarity] - Es

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001

C:Accession: C85990

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
llier, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: AB5480; MUID:21074935; PMID:11206551

A;Accession: C85990

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-156 <STO>

A;Cross-references: GB:AE005174; NID:g12517882; PIDN:AAG58383.1; GSPDB:GN00145; UWGP:

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:
A;Gene: accB

C;Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology

C;Keywords: ligase

Query Match 100.0%; Score 20; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||
DB 49 AAPM 52

RESULT 38

G91144

acetyl-CoA carboxylase (EC 6.4.1.2) biotin carboxyl carrier protein [similarity] - Esche
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 27-Nov-2001
C:Accession: G91144
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: G91144
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <HAY>
A:Cross-references: GB:BA000007; PIDN:BAB37550.1; PID:g13363600; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs4127
C:Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology
C:Keywords: ligase

Query Match 100.0%; Score 20; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||
DB 49 AAPM 52

RESULT 39

AH0912

biotin carboxyl carrier protein [imported] - Salmonella enterica subsp. enterica serovar
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AH0912
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AH0912
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD07894.1; PID:g16504441; GSPDB:GN00176
C:Genetics:
A:Gene: SRY3559
C:Superfamily: biotin carboxyl carrier protein; lipoyl/biotin-binding homology

Query Match 100.0%; Score 20; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||
DB 49 AAPM 52

RESULT 40

D87658

hypothetical protein CC3302 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: D87658
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87658

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <STO>
A:Cross-references: GB:AE005673; NID:g13424996; PIDN:AAK25264.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3302

Query Match 100.0%; Score 20; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
||||
DB 136 AAPM 139

Search completed: December 6, 2002, 13:31:02
Job time : 11.2222 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:20:29 ; Search time 5.55556 Seconds
(without alignments)
29.863 Million cell updates/sec

Title: US-10-033-526-2
Perfect score: 20
Sequence: 1 AAPM 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------|---------------------|
| 1 | 20 | 100.0 | 19 | 1 HHP_THICU | P80487 thibacillu |
| 2 | 20 | 100.0 | 60 | 1 CXI_NAJMO | P01467 naja mossam |
| 3 | 20 | 100.0 | 60 | 1 CXI_NAJPA | P01468 naja pallid |
| 4 | 20 | 100.0 | 71 | 1 GRO1_RABIT | P30782 oryctolagus |
| 5 | 20 | 100.0 | 98 | 1 GROG_BOVIN | O46675 bos taurus |
| 6 | 20 | 100.0 | 103 | 1 GRO_SHEEP | O46678 bos taurus |
| 7 | 20 | 100.0 | 116 | 1 DIUX_DROME | O9YLK4 drosophila |
| 8 | 20 | 100.0 | 121 | 1 YAF1_CAEEL | P52880 caenorhabdi |
| 9 | 20 | 100.0 | 123 | 1 VST1_HEYBU | P29325 hepatitis e |
| 10 | 20 | 100.0 | 123 | 1 VST1_HEYME | Q003499 hepatitis e |
| 11 | 20 | 100.0 | 123 | 1 VST1_HEYME | Q003499 hepatitis e |
| 12 | 20 | 100.0 | 133 | 1 ACPS_DEIRA | O9YXR0 deinococcus |
| 13 | 20 | 100.0 | 142 | 1 YECG_ECOLI | P46888 escherichia |
| 14 | 20 | 100.0 | 142 | 1 YECG_ECOLI | O9YXR0 escherichia |
| 15 | 20 | 100.0 | 146 | 1 VEGA_SHEEP | P50412 ovis aries |
| 16 | 20 | 100.0 | 154 | 1 NABP_RHOSH | O53177 rhodobacter |
| 17 | 20 | 100.0 | 156 | 1 BCCP_ECOLI | P02905 escherichia |
| 18 | 20 | 100.0 | 159 | 1 CYPC_STRCO | P23154 streptomyc |
| 19 | 20 | 100.0 | 161 | 1 BFR_RHOGA | O59738 rhodobacter |
| 20 | 20 | 100.0 | 170 | 1 RL10_CHLDPN | Q929a2 chlamydia p |
| 21 | 20 | 100.0 | 172 | 1 RL10_CHLDPN | Q929a2 chlamydia p |
| 22 | 20 | 100.0 | 172 | 1 RL10_CHLDPN | Q929a2 chlamydia p |
| 23 | 20 | 100.0 | 190 | 1 VEGA_BOVIN | O84319 chlamydia t |
| 24 | 20 | 100.0 | 190 | 1 VEGA_BOVIN | P15691 bos taurus |
| 25 | 20 | 100.0 | 190 | 1 VEGA_PIG | O9QKR0 equus caball |
| 26 | 20 | 100.0 | 191 | 1 CBP2_ARATH | P49151 sus scrofa |
| 27 | 20 | 100.0 | 198 | 1 ACD2_CLOPE | P30187 arabidopsis |
| 28 | 20 | 100.0 | 200 | 1 ACPD_ECO57 | O8X9S9 escherichia |
| 29 | 20 | 100.0 | 200 | 1 ACPD_ECO57 | O8X9S9 escherichia |
| 30 | 20 | 100.0 | 200 | 1 ACPD_ECO57 | O8X9S9 escherichia |
| 31 | 20 | 100.0 | 200 | 1 ACPD_ECO57 | O8X9S9 escherichia |
| 32 | 20 | 100.0 | 213 | 1 ACPD_PSEAE | Q8E60 yersinia pe |
| 33 | 20 | 100.0 | 214 | 1 VEGA_CANRA | O9H17 pseudomonas |
| | | | | | Q9myv3 canis famill |

| | | | | | |
|----|----|-------|-----|---------------|--------------------|
| 34 | 20 | 100.0 | 228 | 1 RS5_DROME | Q24186 drosophila |
| 35 | 20 | 100.0 | 229 | 1 VMAT_VSVSJ | P03519 vesicular s |
| 36 | 20 | 100.0 | 229 | 1 YXH_K_CAEEL | Q20728 caenorhabdi |
| 37 | 20 | 100.0 | 231 | 1 YR01_CAEEL | Q10014 caenorhabdi |
| 38 | 20 | 100.0 | 232 | 1 VEGA_HUMAN | P15692 homo sapien |
| 39 | 20 | 100.0 | 232 | 1 YR26_MYCTU | O50702 mycobacteri |
| 40 | 20 | 100.0 | 237 | 1 VMAT_VSVIG | P04876 vesicular s |
| 41 | 20 | 100.0 | 240 | 1 SESA_AERPE | O9YEX9 aeropyrum p |
| 42 | 20 | 100.0 | 246 | 1 BDNF_CHICK | P25429 gallus gall |
| 43 | 20 | 100.0 | 247 | 1 BDNF_FELCA | O9LST3 felis silve |
| 44 | 20 | 100.0 | 247 | 1 BDNF_HUMAN | P23560 homo sapien |
| 45 | 20 | 100.0 | 247 | 1 BDNF_PROLO | O18755 procyon lot |

ALIGNMENTS

| | | | | | |
|---|-----------|----------|-------------------------|--|--|
| RESULT 1 | | | | | |
| HHP_THICU | | | | | |
| ID HHP_THICU | STANDARD; | PRT; | 19 AA. | | |
| AC P80487; | | | | | |
| DT 01-OCT-1996 (Rel. 34, Created) | | | | | |
| DT 01-OCT-1996 (Rel. 34, Last sequence update) | | | | | |
| DT 01-OCT-1996 (Rel. 34, Last annotation update) | | | | | |
| DE Heterotroph-specific protein (Fragment). | | | | | |
| OS Thibacillus cuprinus. | | | | | |
| OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Thiomonas. | | | | | |
| OX NCBI_TaxID=36860; | | | | | |
| RN [1] | | | | | |
| RP SEQUENCE. | | | | | |
| RC STRAIN=DSM 5494; | | | | | |
| RA Marin I., Amaro A.M., Jerez C.A., Amils R., Abad J.P.; | | | | | |
| RL Submitted (SEP-1995) to the SWISS-PROT data bank. | | | | | |
| CC -I- MISCELLANEOUS. FOUND SPECIFICALLY IN CELLS CULTURED | | | | | |
| CC HERETROTROPHICALLY. | | | | | |
| FT NON_TER | 19 | 19 | | | |
| SQ SEQUENCE | 19 AA; | 1786 MW; | C549197D0A492B07 CRC64; | | |

Query Match 100.0%; Score 20; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
DB 12 AAPM 15

RESULT 2
CXI_NAJMO STANDARD; PRT; 60 AA.
ID CXI_NAJMO
AC P01467;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytotoxin 1 (Cardiotoxin XIIB) (Cytotoxin V-II-1) (CTX IIB).
OS Naja mossambica (Mozambique cobra).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Naja.
OX NCBI_TaxID=8644;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RA Louw A.I.;
RT "Snake venom toxins. The amino acid sequences of three cytotoxin
homologues from Naja mossambica mossambica venom.";
RL Biochim. Biophys. Acta 336:481-495(1974).
RN [2]
RP STRUCTURE BY NMR.
RA MEDLINE=88029481; PubMed=2822421;
RX Oetting G., Steinmetz W.E., Bougis P.E., Rochat H., Wuethrich K.;
RT "Sequence-specific 1H-NMR assignments and determination of the
secondary structure in aqueous solution of the cardiotoxins CTXIa

```
RT and CTXIIb from Naja mossambica mossambica.";
RL Eur. J. Biochem. 168:609-620(1987).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=93279339; PubMed=8504828;
RA O'Connell J.F., Bougis P.E., Wuehrich K.;
RT "Determination of the nuclear-magnetic-resonance solution structure
of cardiotoxin CTx IIB from Naja mossambica mossambica.";
RL Eur. J. Biochem. 213:891-900(1993).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- MISCELLANEOUS: LD(50) IS 0.83 MG/KG BY INTRAVENOUS INJECTION.
CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
DR PIR; A01728; H3NJ1M.
DR PIR; S02517; S02517.
DR PDB; 2CCX; 31-JAN-94.
DR InterPro; IPR003572; Cytotoxin.
DR InterPro; IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR PRINTS; PR00282; CYTOTOXIN.
DR PRODOM; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Venom; Cytotoxin; Cardiotoxin; Multigene family; 3D-structure.
FT DISULFID 3 21
FT DISULFID 14 38
FT DISULFID 42 53
FT DISULFID 54 59
FT STRAND 2 3
FT TURN 9 10
FT STRAND 12 13
FT TURN 16 17
FT STRAND 21 26
FT TURN 30 31
FT STRAND 35 37
FT STRAND 49 54
FT TURN 57 58
SQ SEQUENCE 60 AA; 6826 MW; 68BE50B776B6491C CRC64;

Query Match 100.0%; Score 20; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
    ||||
Db 28 AAPM 31

RESULT 3
CX1_NAJPA
ID CX1_NAJPA STANDARD; PRT; 60 AA.
AC P01468;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytotoxin 1 (Cardiotoxin gamma).
OS Naja pallida (Red spitting cobra).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Naja.
OX NCBI_TaxID=8658;
RN [1]
RP SEQUENCE, AND DISULFIDE BONDS.
RC TISSUE=Venom;
RX MEDLINE=75205552; PubMed=1148181;
RA Fryklund L., Eaker D.;
RT "The complete covalent structure of a cardiotoxin from the venom of
Naja nigricollis (African black-necked spitting cobra).";
RL Biochemistry 14:2865-2871(1975).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.55 ANGSTROMS).
RX MEDLINE=94254084; PubMed=8196041;
RA Bilwes A., Rees B., Moras D., Menez R., Menez A.;
RT "X-ray structure at 1.55 A of toxin gamma, a cardiotoxin from Naja
```

```
RT nigricollis venom. Crystal packing reveals a model for insertion into
RT membranes.";
RL J. Mol. Biol. 239:122-136(1994).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
CC -1- CAUTION: THE VENOM OF THIS SNAKE WAS ORIGINALLY THOUGHT TO BE THAT
OF N.NIGRICOLLIS WHILE IT IS REALLY FROM N.PALLIDA.
DR PIR; A37578; H3NJ1B.
DR PDB; 1TGX; 30-APR-94.
DR PDB; 1CXN; 20-DEC-94.
DR PDB; 1CXO; 20-DEC-94.
DR InterPro; IPR003572; Cytotoxin.
DR InterPro; IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR PRINTS; PR00282; CYTOTOXIN.
DR PRODOM; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Venom; Cytotoxin; Cardiotoxin; Multigene family; 3D-structure.
FT DISULFID 3 21
FT DISULFID 14 38
FT DISULFID 42 53
FT DISULFID 54 59
FT STRAND 2 4
FT TURN 8 9
FT STRAND 11 13
FT TURN 16 17
FT STRAND 20 26
FT TURN 27 28
FT STRAND 29 29
FT TURN 30 31
FT STRAND 35 39
FT STRAND 49 54
FT TURN 57 58
SQ SEQUENCE 60 AA; 6827 MW; 68BE50B996B6491C CRC64;
```

```
Query Match 100.0%; Score 20; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAPM 4
    ||||
Db 28 AAPM 31

RESULT 4
GRO1_RABIT
ID GRO1_RABIT STANDARD; PRT; 71 AA.
AC P30782;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE permeability factor 2 (RPF2) (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Alveolar macrophage;
RX MEDLINE=95129889; PubMed=7828903;
RA Johnson M.C., Goodman R.B. II, Kajikawa O., Wong V.A., Mongovin S.M.,
RA Martin T.R.;
RT "Cloning of two rabbit GRO homologues and their expression in
alveolar macrophages.";
RL Gene 151:337-338(1994).
RN [2]
RP SEQUENCE OF 1-36.
RC STRAIN=New Zealand white; TISSUE=Peritoneal cavity;
RX MEDLINE=91378900; PubMed=1898341;
RA Jose P.J., Collins P.D., Perkins J.A., Beaubien B.C., Totty N.F.,
RA Waterfield M.D., Hsuan J., Williams T.J.;
RT "Identification of a second neutrophil-chemoattractant cytokine
```

```

RT generated during an inflammatory reaction in the rabbit peritoneal
RT cavity in vivo. Purification, partial amino acid sequence and
RT structural relationship to melanoma-growth-stimulatory activity.";
RL Biochem. J. 278:493-497(1991).
CC -!- FUNCTION: HAS CHEMOTACTIC ACTIVITY FOR NEUTROPHILS (BY
CC SIMILARITY).
CC -!- SUBUNIT: HOMODIMER (PROBABLE).
CC -!- INDUCTION: GENERATED DURING AN INFLAMMATORY REACTION.
CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC C-X-C) (CHEMOKINE CXC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L19157; AAA64357.1; -.
DR PIR: S17507; S17507.
DR HSSP: P19875; 1QNK.
DR InterPro: IPR001089; CXC_chmkine_sm11.
DR InterPro: IPR001811; Chemokine_IL8.
DR Pfam: PF00048; IL8; 1.
DR SMART: SM00199; SCY; 1.
DR PROSITE: PS00471; SMALL_CYTOKINES_CXC; PARTIAL.
KW Cytokine; Growth factor; Inflammatory response.
FT DISULFID 7 33 BY SIMILARITY.
FT DISULFID 9 49 BY SIMILARITY.
FT CONFLICT 20 20 N -> S (IN REF. 2).
FT CONFLICT 23 23 N -> S (IN REF. 2).
FT NON_TER 71 71
SQ SEQUENCE 71 AA; 7713 MW; E91387CB382C8008 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 71;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 53 AAPM 56

RESULT 5
GROG_BOVIN STANDARD; PRT; 98 AA.
ID GROG_BOVIN
AC 046675;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Growth regulated protein homolog gamma precursor (GRO-gamma).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99152612; PubMed=10028286;
RA Modi W.S., Yoshimura T.;
RT "Isolation of novel GRO genes and a phylogenetic analysis of the CXC
RT chemokine subfamily in mammals.";
RL Mol. Biol. Evol. 16:180-193(1999).
CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC C-X-C) (CHEMOKINE CXC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U95811; AAB93927.1; -.
DR HSSP: P19875; 1QNK.
DR InterPro: IPR001089; CXC_chmkine_sm11.
DR InterPro: IPR001811; Chemokine_IL8.
DR Pfam: PF00048; IL8; 1.
DR PRINTS: PR00437; SMALLCYTKCXC.
DR SMART: SM00199; SCY; 1.
DR PROSITE: PS00471; SMALL_CYTOKINES_CXC; 1.
KW Cytokine; Growth factor; Inflammatory response; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 98 GROWTH REGULATED PROTEIN HOMOLOG GAMMA.
FT DISULFID 39 65 BY SIMILARITY.
FT DISULFID 41 81 BY SIMILARITY.
SQ SEQUENCE 98 AA; 10393 MW; 942CD6897C21EDE9 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 85 AAPM 88

RESULT 6
GRO_SHEEP STANDARD; PRT; 103 AA.
ID GRO_SHEEP
AC 046678;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Growth regulated protein precursor (CXCL1).
GN SCYB1 OR GRO.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99152612; PubMed=10028286;
RA Modi W.S., Yoshimura T.;
RT "Isolation of novel GRO genes and a phylogenetic analysis of the CXC
RT chemokine subfamily in mammals.";
RL Mol. Biol. Evol. 16:180-193(1999).
CC -!- FUNCTION: HAS CHEMOTACTIC ACTIVITY FOR NEUTROPHILS.
CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC C-X-C) (CHEMOKINE CXC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U95814; AAB93930.1; -.
DR HSSP: P19875; 1QNK.
DR InterPro: IPR001089; CXC_chmkine_sm11.
DR InterPro: IPR001811; Chemokine_IL8.
DR Pfam: PF00048; IL8; 1.
DR PRINTS: PR00437; SMALLCYTKCXC.
DR SMART: SM00199; SCY; 1.
DR PROSITE: PS00471; SMALL_CYTOKINES_CXC; 1.
KW Cytokine; Growth factor; Inflammatory response; Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 103 GROWTH REGULATED PROTEIN.
FT DISULFID 39 65 BY SIMILARITY.
FT DISULFID 41 81 BY SIMILARITY.
SQ SEQUENCE 103 AA; 10820 MW; C59857F346716695 CRC64;

```



```

Query Match          100.0%; Score 20; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
      1111
Db      85 AAPM 88

RESULT 7
DIUX_DROME
ID DIUX_DROME STANDARD; PRT; 116 AA.
AC Q9VLK4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Diuretic hormone class-II precursor (Diuretic peptide) (DP) (DH(31)).
GN CG13094.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-T., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP FUNCTION, AND SYNTHESIS OF 76-106.
RX MEDLINE=21216780; Pubmed=11316500;
RA Coast G.M., Webster S.G., Scheggs K.M., Tobe S.S., Schooley D.A.;
RT "The Drosophila melanogaster homologue of an insect calcitonin-like
RT diuretic peptide stimulates V-ATPase activity in fruit fly Malpighian
RT tubules.";
```

```

RL J. Exp. Biol. 204:1795-1804(2001).
CC -I- FUNCTION: REGULATION OF FLUID SECRETION. STIMULATES MALPIGHIAN
CC TUBULES FLUID SECRETION BY ACTIVATING THE APICAL MEMBRANE V-ATPASE
CC VIA CYCLIC AMP OF PRINCIPAL CELLS IN THE MAIN SECRETORY SEGMENT.
CC -I- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -I- SIMILARITY: BELONGS TO THE DIURETIC HORMONE CLASS II FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE003621; AAF52685.1; -.
CC FlyBase; FBgn0032048; CG13094.
KW Hormone; Amidation; Cleavage on pair of basic residues; Signal.
FT SIGNAL 1 25
FT PROPEP 26 75
FT PEPTIDE 76 106
FT PROPEP 112 116
FT MOD_RES 106 106
FT AMIDATION (G-107 PROVIDE AMIDE GROUP)
FT (POTENTIAL).
SQ SEQUENCE 116 AA; 12622 MW; 1BE68A1EA8FA49 CRC64;
```

```

Query Match          100.0%; Score 20; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 AAPM 4
      1111
Db      25 AAPM 28
```

```

RESULT 8
YAF1_CAEEL
ID YAF1_CAEEL STANDARD; PRT; 121 AA.
AC P52880;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 13.4 kDa protein F46C5.1 in chromosome II.
GN F46C5.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Thomas K.;
RL Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z54281; CAA91048.1; -.
CC WormPep; F46C5.1; CE03343.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 6 26
FT SEQUENCE 121 AA; 13440 MW; 477FD9C1771CDF01 CRC64;
```

```

Query Match          100.0%; Score 20; DB 1; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPM 4
```

```
Db      22 AAPM 25

      ||||
RESULT 9
VSTL_HEVBU      STANDARD;      PRT;      123 AA.
ID VSTL_HEVBU
AC P29325;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 1 (ORF3).
OS Hepatitis E virus (strain Burma) (HEV), and
OS Hepatitis E virus (strain Pakistan) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31767, 33774;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Burma;
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Pakistan;
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
RN [3]
RP SEQUENCE FROM N.A.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M73218; AAA45735.1; -.
CC EMBL: M80581; AAA45726.1; -.
DR PIR: B40778; VHMWME.
DR InterPro: IPR003384; HEV_ORF2.
DR Pfam: PF02444; HEV_ORF2; 1.
SQ SEQUENCE 123 AA; 12676 MW; 8A5A798B1B74EDE5 CRC64;

Query Match      100.0%; Score 20; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
      ||||
Db      7 AAPM 10

RESULT 10
VSTL_HEVME      STANDARD;      PRT;      123 AA.
ID VSTL_HEVME
AC Q03499;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Structural protein 1.
OS Hepatitis E virus (strain Mexico) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31768;
RN [1]
RP SEQUENCE FROM N.A.

Rx MEDLINE=93079857; PubMed=1448913;
Ra Huang C.C., Nguyen D., Fernandez J., Yun K.Y., Fry K.E.,
Ra Bradley D.W., Tam A.W., Reyes G.R.;
Rt "Molecular cloning and sequencing of the Mexico isolate of hepatitis
Rt E virus (HEV).";
Rl Virology 191:550-558(1992).
Cc -----
Cc This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc or send an email to license@isb-sib.ch).
Cc -----
Cc EMBL: M74506; AAA45731.1; -.
DR PIR: C44212; C44212.
DR InterPro: IPR003384; HEV_ORF2.
DR Pfam: PF02444; HEV_ORF2; 1.
SQ SEQUENCE 123 AA; 12714 MW; C888F5D638852A68 CRC64;

Query Match      100.0%; Score 20; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
      ||||
Db      7 AAPM 10

RESULT 11
VSTL_HEVMY      STANDARD;      PRT;      123 AA.
ID VSTL_HEVMY
AC Q04612;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Structural protein 1 (ORF3).
OS Hepatitis E virus (strain Myanmar) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227573; PubMed=8470371;
RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Rikihisa T., Winn K.;
RT "Sequence and gene structure of the hepatitis E virus isolated from
RT Myanmar.";
Rl Virus Genes 7:95-109(1993).
Cc -----
Cc This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc or send an email to license@isb-sib.ch).
Cc -----
Cc EMBL: D10330; BAA01173.1; -.
DR InterPro: IPR003384; HEV_ORF2.
DR Pfam: PF02444; HEV_ORF2; 1.
SQ SEQUENCE 123 AA; 12751 MW; FB81143F0B31F8A4 CRC64;

Query Match      100.0%; Score 20; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
      ||||
Db      7 AAPM 10
```

```
RESULT 12
ACPS_DEIRA          STANDARD;          PRT;          133 AA.
ID  ACPS_DEIRA
AC  Q9RXR0;
DT  15-JUN-2002 (Rel. 41, Created)
DT  15-JUN-2002 (Rel. 41, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Holo-[acyl-carrier protein] synthase (EC 2.7.8.7) (Holo-ACP synthase)
DE  (4'-phosphopantetheinyl transferase acps).
GN  ACPS OR DR0247.
OS  Deinococcus radiodurans.
OC  Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC  Deinococcaceae; Deinococcus.
OX  NCBI_TaxID=1299;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=R1;
RX  MEDLINE=20036896; PubMed=10567266;
RA  White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA  Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
RA  Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA  Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA  Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA  Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA  Fraser C.M.;
RT  "Genome sequence of the radioresistant bacterium Deinococcus
RT  radiodurans R1.";
RL  Science 286:1571-1577(1999).
CC  -1- FUNCTION: Transfers the 4'-phosphopantetheine moiety from coenzyme
CC  A to a Ser of acyl-carrier protein (By similarity).
CC  -1- CATALYTIC ACTIVITY: CoA + apo-[acyl-carrier protein] = adenosine
CC  3',5'-bisphosphate + holo-[acyl-carrier protein].
CC  -1- COFACTOR: Magnesium (By similarity).
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC  -1- SIMILARITY: BELONGS TO THE P-PANT TRANSFERASE SUPERFAMILY. ACPS
CC  FAMILY.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
CC  EMBL; AE001885; AAF09829.1; -.
DR  TIGR; DR0247; -.
DR  InterPro; IPR002582; ACPS.
DR  InterPro; IPR004568; Pantethn_trn.
DR  Pfam; PF01648; ACPS; 1.
DR  ProDom; PD004282; ACPS; 1.
DR  TIGRFAMs; TIGR00556; pantethn_trn; 1.
KW  Transferase; Lipid synthesis; Fatty acid biosynthesis; Magnesium;
KW  Complete proteome.
FT  METAL 8
FT  METAL 56 MAGNESIUM (BY SIMILARITY).
SQ  SEQUENCE 133 AA; 15114 MW; C5E0009B9728D120 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 86 AAPM 89

RESULT 13
YECG_ECOLI
ID  YECG_ECOLI STANDARD; PRT; 142 AA.
AC  P46888;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
OS  Salmone
```

```
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Hypothetical protein yecG.
GN  YECG OR B1895.
OS  Escherichia coli.
OC  Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC  Escherichia.
OX  NCBI_TaxID=562;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=K12 / EMG2;
RA  Estep P., O'Keefe T., Robison K., Church G.M.;
RL  Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=K12 / MG1655;
RX  MEDLINE=97426617; PubMed=9278503;
RA  Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA  Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA  Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA  Mau B., Shao Y.;
RT  "The complete genome sequence of Escherichia coli K-12.";
RL  Science 277:1453-1474(1997).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN=K12;
RX  MEDLINE=97251358; PubMed=9097040;
RA  Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA  Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA  Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA  Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA  Saito N., Sempel G., Seki Y., Sivasubaram S., Tagami H.,
RA  Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT  "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT  corresponding to the 40.1-50.0 min region on the linkage map.";
RL  DNA Res. 3:379-392(1996).
RN  -1- SIMILARITY: BELONGS TO THE UPF0022 (USPA) FAMILY.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
CC  EMBL; U27211; AAA68603.1; -.
DR  EMBL; AE000283; AAC74965.1; -.
DR  EMBL; D90831; BAA15716.1; -.
DR  EcoGene; EG12862; yecG.
DR  InterPro; IPR000041; Usp.
DR  Pfam; PF00582; usp; 1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 142 AA; 15633 MW; 87DFEE19A8EE58FC CRC64;

Query Match 100.0%; Score 20; DB 1; Length 142;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 50 AAPM 53

RESULT 14
YECG_SALTY
ID  YECG_SALTY STANDARD; PRT; 142 AA.
AC  Q9RM66;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Hypothetical protein yecG.
GN  YECG OR USPS OR STM1927.
OS  Salmone
```

```
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20052966; PubMed=10586519;
RA Yanagihara S., Iyoda S., Ohnishi K., Iino T., Kutsukake K.;
RT "Structure and transcriptional control of the flagellar master operon
of Salmonella typhimurium.";
RL Genes Genet. Syst. 74:105-111(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
CC -!- SIMILARITY: BELONGS TO THE UPF0022 (USPA) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D43640; BAA85313.1; -.
CC EMBL: AE008786; AAL20843.1; -.
CC StyGene; SG??777; yecg.
CC InterPro: IPR000041; Usp.
CC Pfam: PF00582; Usp; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 142 AA; 15498 MW; F39A8EE5B1777C2 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 142;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
   ||||
Db 50 AAPM 53

RESULT 15
VEGA_SHEEP STANDARD; PRT; 146 AA.
ID VEGA_SHEEP
AC P50412;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
DE permeability factor) (VPF).
GN VEGF OR VEGFA.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97117958; PubMed=8958842;
RA Redmer D.A., Dai Y., Li J., Charnock-Jones D.S., Smith S.K.,
RA Reynolds L.P., Moor R.M.;
RT "Characterization and expression of vascular endothelial growth
RT factor (VEGF) in the ovine corpus luteum.";
RL J. Reprod. Fertil. 108:157-165(1996).
```

```
CC -!- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
CC endothelial cell growth. It induces endothelial cell
CC proliferation, promotes cell migration, inhibits apoptosis, and
CC induces permeabilization of blood vessels. It binds to the
CC VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
CC heparin (By similarity).
CC -!- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
CC with PLGF (By similarity).
CC -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X89506; CAA61677.1; -.
CC HSSP: P15692; IVP.
CC InterPro: IPR000072; PD_growth_factor.
CC Pfam: PF00341; PDGF; 1.
CC ProDom: PD001629; PD_growth_factor; 1.
CC SMART: SM00141; PDGF; 1.
CC PROSITE: PS00249; PDGF_1; 1.
CC PROSITE: PS50278; PDGF_2; 1.
KW Mitogen; Angiogenesis; Growth factor; Glycoprotein; Signal;
KW Heparin-binding; Multigene family.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 146 VASCULAR ENDOTHELIAL GROWTH FACTOR A.
FT DISULFID 51 93 BY SIMILARITY.
FT DISULFID 82 127 BY SIMILARITY.
FT DISULFID 86 129 BY SIMILARITY.
FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 146 AA; 17247 MW; 4E792CB557F91760 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 146;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
   ||||
Db 26 AAPM 29

RESULT 16
NAPB_RHOSH STANDARD; PRT; 154 AA.
ID NAPB_RHOSH
AC Q53177;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Diheme cytochrome c napb precursor.
GN NAPB.
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RX MEDLINE=96332666; PubMed=8730872;
RA Reyes F., Roldan M.D., Klipp W., Castillo F., Moreno-Vivian C.;
RT "Isolation of periplasmic nitrate reductase genes from Rhodobacter
RT sphaeroides DSM 158: structural and functional differences among
RT prokaryotic nitrate reductases.";
RL Mol. Microbiol. 19:1307-1318(1996).
CC -!- FUNCTION: SMALL SUBUNIT OF THE PERIPLASMIC NITRATE REDUCTASE
CC (NAP). ONLY EXPRESSED AT HIGH LEVELS DURING AEROBIC GROWTH. NAPAB
CC COMPLEX RECEIVES ELECTRONS FROM THE MEMBRANE-ANCHORED TETRAHEME
CC NAPC PROTEIN, THUS ALLOWING ELECTRON FLOW BETWEEN MEMBRANE AND
```



```
CC PERIPLASM, ESSENTIAL FUNCTION FOR NITRATE ASSIMILATION AND MAY
CC HAVE A ROLE IN ANAEROBIC METABOLISM.
CC -1- SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT AND A CYTOCHROME C (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- INDUCTION: INDUCED BY NITRATE. NOT REPRESSED BY AMMONIUM OR
CC OXYGEN.
CC -1- PTM: BINDS TWO HEME GROUPS PER MOLECULE (BY SIMILARITY).
CC -1- SIMILARITY: STRONG, TO OTHER BACTERIAL NABP.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 246806; -; NOT_ANNOTATED_CDS.
CC DR InterPro; IPR000345; CytC_heme_bind.
CC DR PROSITE; PS00190; CYTOCHROME_C; 2.
CC KW Electron transport; Heme; Periplasmic; Signal.
CC FT SIGNAL 1 24 POTENTIAL.
CC FT CHAIN 25 154 DIHEME CYTOCHROME C NABP.
CC FT MOD_RES 25 25 PYRROLIDONE CARBOXYLIC ACID (POTENTIAL).
CC FT BINDING 82 82 HEME 1 (COVALENT) (BY SIMILARITY).
CC FT BINDING 85 85 HEME 1 (COVALENT) (BY SIMILARITY).
CC FT METAL 86 86 IRON 1 (HEME AXIAL LIGAND)
CC (BY SIMILARITY).
CC FT BINDING 122 122 HEME 2 (COVALENT) (BY SIMILARITY).
CC FT BINDING 125 125 HEME 2 (COVALENT) (BY SIMILARITY).
CC FT METAL 126 126 IRON 2 (HEME AXIAL LIGAND)
CC (BY SIMILARITY).
CC SQ SEQUENCE 154 AA; 16908 MW; 63CE7404A5864977 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 95 AAPM 98

RESULT 17
BCCP_ECOLI STANDARD; PRT; 156 AA.
ID BCCP_ECOLI
AC P02905;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP).
GN ACCB OR FABE OR B3255 OR Z4615 OR ECS4127.
OS Escherichia coli, and
OS Escherichia coli 0157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=89282408; PubMed=2660106;
RA Muramatsu S., Mizuno T.;
RT "Nucleotide sequence of the fabE gene and flanking regions containing
RL a bent DNA sequence of Escherichia coli.";
RL Nucleic Acids Res. 17:3982-3982(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92112819; PubMed=1370469;
RA Li S.-J., Cronan J.E. Jr.;
RT "The gene encoding the biotin carboxylase subunit of Escherichia coli
RL acetyl-CoA carboxylase.";
RL J. Biol. Chem. 267:855-863(1992).
```

```
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90126231; PubMed=2575489;
RA Alix J.-H.;
RT "A rapid procedure for cloning genes from lambda libraries by
RT complementation of E. coli defective mutants: application to the fabE
RT region of the E. coli chromosome.";
RL DNA 8:779-789(1989).
RN [4]
RP SEQUENCE FROM N.A.
RA Best E.A., Knauf V.C.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grodbeck E.J., Davis N.W., Llm A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RL O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [8]
RP SEQUENCE OF 1-25 FROM N.A.
RX MEDLINE=93123150; PubMed=7678242;
RA Li S.-J., Cronan J.E. Jr.;
RT "Growth rate regulation of Escherichia coli acetyl coenzyme A
RT carboxylase, which catalyzes the first committed step of lipid
RT biosynthesis.";
RL J. Bacteriol. 175:332-340(1993).
RN [9]
RP SEQUENCE OF 75-156.
RX MEDLINE=77187896; PubMed=324999;
RA Sutton M.R., Fall R.R., Nervi A.M., Alberts A.W., Vagelos P.R.,
RA Bradshaw R.A.;
RT "Amino acid sequence of Escherichia coli biotin carboxyl carrier
RT protein (9100).";
RL J. Biol. Chem. 252:3934-3940(1977).
RN [10]
RP SEQUENCE OF 133-156 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92052166; PubMed=1682920;
RA Kondo H., Shiratsuchi K., Yoshimoto T., Masuda T., Kitazono A.,
RA Tsuru D., Anai M., Sekiguchi M., Tanabe T.;
RT "Acetyl-CoA carboxylase from Escherichia coli: gene organization and
RL nucleotide sequence of the biotin carboxylase subunit.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9730-9733(1991).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 77-156.
RX MEDLINE=96363677; PubMed=8747466;
```



```
RA Athappilly F.K., Hendrickson W.A.;
RT "Structure of the biotinyl domain of acetyl-coenzyme A carboxylase
RT determined by MAD phasing.";
RL Structure 3:1407-1419(1995).
RN [12]
RP STRUCTURE BY NMR OF 70-156.
RX MEDLINE=98060761; PubMed=9398236;
RA Yao X., Wei D., Soden C. Jr., Summers M.F., Beckett D.;
RT "Structure of the carboxy-terminal fragment of the apo-biotin
RT carboxyl carrier subunit of Escherichia coli acetyl-CoA
RL carboxylase.";
RL Biochemistry 36:15089-15100(1997).
RN [13]
RP STRUCTURE BY NMR OF 77-156.
RX MEDLINE=99230195; PubMed=10213607;
RA Roberts E.L., Shu N., Howard M.J., Broadhurst R.W., Chapman-Smith A.,
RA Wallace J.C., Morris T., Cronan J.E. Jr., Perham R.N.;
RT "Solution structures of apo and holo biotinyl domains from acetyl
RT coenzyme A carboxylase of Escherichia coli determined by triple-
RT resonance nuclear magnetic resonance spectroscopy.";
RL Biochemistry 38:5045-5053(1999).
CC -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE ACETYL COENZYME A
CC CARBOXYLASE COMPLEX; FIRST, BIOTIN CARBOXYLASE CATALYZES THE
CC TRANSFERS THE CARBOXYL GROUP TO FORM MALONYL-COA.
CC -!- PATHWAY: long-chain fatty acid biosynthesis; first step.
CC -!- SUBUNIT: HOMODIMER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14825; CAA32933.1; -.
DR EMBL: M80458; AAA23408.1; -.
DR EMBL: M32214; AAA23744.1; -.
DR EMBL: M79446; -; NOT_ANNOTATED_CDS.
DR EMBL: M83198; AAA23745.1; -.
DR EMBL: U18997; AAA58058.1; -.
DR EMBL: AE000404; AAC76287.1; -.
DR EMBL: AE005553; AAG58383.1; -.
DR EMBL: AP002564; BAB37550.1; -.
DR EMBL: S52932; AAB24892.1; -.
DR PIR: A03402; BKEC9.
DR PIR: A33643; A33643.
DR PIR: JS0686; JS0686.
DR PIR: PS0272; PS0272.
DR PIR: C40637; C40637.
DR PDB: 1BDO; 01-AUG-96.
DR PDB: 2BDO; 27-APR-99.
DR PDB: 3BDO; 26-APR-99.
DR PDB: 1A6X; 14-OCT-98.
DR SWISS-2DPAGE: P02905; COLI.
DR EcoGene; EG10275; accB.
DR InterPro; IPR001249; AccCoA_biotinCC.
DR InterPro; IPR001882; Biotin_attach.
DR InterPro; IPR000089; Biotin_lipoyl.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR PRINTS; PR01071; ACOABIOITNCC.
DR TIGRFAMS; TIGR00531; BCCP; 1.
DR PROSITE; PS00188; BIOTIN; 1.
KW Fatty acid biosynthesis; Biotin; 3D-structure; Complete proteome.
FT BINDING 122 122 BIOTIN.
FT CONFLICT 113 113 D -> N (IN REF. 3).
SQ SEQUENCE 156 AA; 16687 MW; 05FFDCB912A683A3 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
DB 49 AAPM 52

RESULT 18
CYPC_STRCO
ID CYPC_STRCO STANDARD; PRT; 159 AA.
AC P23154;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE putative polyketide cyclase (White ORF VI).
GN SC05315 OR SC669.18.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=91171868; PubMed=2077356;
RA Davis N.K., Chater K.F., Bruton C.J.;
RT "Spore colour in Streptomyces coelicolor A3(2) involves the
RT developmentally regulated synthesis of a compound biosynthetically
RT related to polyketide antibiotics.";
RL Mol. Microbiol. 4:1679-1691(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- FUNCTION: INVOLVED IN DEVELOPMENTALLY REGULATED SYNTHESIS OF A
CC COMPOUND BIOSYNTHEMICALLY RELATED TO POLYKETIDE ANTIBIOTICS,
CC WHICH IS ESSENTIAL FOR SPORE COLOR IN STREPTOMYCES COELICOLOR.
CC -!- SIMILARITY: TO POLYKETIDE CYCLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X55942; CAA39411.1; -.
DR EMBL: AL079356; CAB45609.1; -.
DR PIR: S11977; S11977.
KW Complete proteome.
SQ SEQUENCE 159 AA; 18197 MW; EBEAD2FCCAB23A43 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
DB 12 AAPM 15

RESULT 19
BFR_RHOCA
```

```

ID   BFR_RHOCA                STANDARD;             PRT;             161 AA.
AC   Q59738;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Bacterioferritin (BFR).
GN   BFR.
OS   Rhodobacter capsulatus (Rhodopseudomonas capsulata).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC   Rhodobacter.
OX   NCBI_TaxID=1061;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 938 / 37b4;
RX   MEDLINE=96275893; PubMed=8674981;
RA   Penfold C.N., Ringeling P.L., Davy S.L., Moore G.R., McEwan A.G.,
RA   Spiro S.;
RT   "Isolation, characterisation and expression of the bacterioferritin
RT   gene of Rhodobacter capsulatus.";
RL   FEMS Microbiol. Lett. 139:143-148(1996).
CC   -!- FUNCTION: May perform analogous functions in iron detoxification
CC   and storage to that of animal ferritins (By similarity).
CC   -!- SUBUNIT: OLIGOMER OF 24 IDENTICAL SUBUNITS (BY SIMILARITY).
CC   -!- MISCELLANEOUS: Bacterioferritin contains protoheme IX in addition
CC   to the non-haem iron core.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIOFERRITIN FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; 254247; CA91017.1; -.
DR   HSSP; P11056; 1BCF.
DR   InterPro; IPR002024; Bacterioferritin.
DR   Pfam; PF01334; Bacteriofer; 1.
DR   PRINTS; PR00601; BACFERRITIN.
DR   ProDom; PD002269; Bacterioferritin; 1.
DR   TIGRfams; TIGR00754; bfr; 1.
DR   PROSITE; PS00549; BACTERIOFERRITIN; 1.
KW   Iron storage; Heme.
FT   METAL 52 52 IRON (HEME AXIAL LIGAND) (POTENTIAL).
SQ   SEQUENCE 161 AA; 18172 MW; 9E534CBC531EC709 CRC64;

Query Match 100.0%; Score 20; DB 1; length 161;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 154 AAPM 157

RESULT 20
RL10_CHLPN STANDARD; PRT; 170 AA.
ID RL10_CHLPN
AC Q9Z9A2; Q9JQB8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L10.
GN RPLJ OR RL10 OR CPN0079 OR CP0696.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammell C., Fan J., Hyman R.W.,

```

```

RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MOPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR   EMBL; AE001593; AAD18232.1; -.
DR   EMBL; AE002228; AAF38504.1; -.
DR   EMBL; AP002545; BAA98289.1; -.
DR   PHCI-2DPAGE; Q9Z9A2; -.
DR   TIGR; CP0696; -.
DR   InterPro; IPR001790; Ribosomal_L10.
DR   InterPro; IPR002363; Ribosomal_L10eub.
DR   Pfam; PF00466; Ribosomal_L10; 1.
DR   PROSITE; PS01109; RIBOSOMAL_L10; FALSE_NEG.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE 170 AA; 18429 MW; 148263C17E062632 CRC64;

Query Match 100.0%; Score 20; DB 1; length 170;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 142 AAPM 145

RESULT 21
RL10_CHLMU STANDARD; PRT; 172 AA.
ID RL10_CHLMU
AC Q9PK78;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L10.
GN RPLJ OR TC0591.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPn / Nigg;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,

```

RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -I- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE002328; AAF39423.1; -.
CC TIGR: TC0591; -.
DR InterPro: IPR001790; Ribosomal_L10.
DR InterPro: IPR002363; Ribosomal_L10eub.
DR Pfam: PF00466; Ribosomal_L10; 1.
DR PROSITE: PS01109; RIBOSOMAL_L10; 1.
KM Ribosomal protein: Complete proteome.
SQ SEQUENCE 172 AA; 18817 MW; 9579AE1D42268483 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 172;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 142 AAPM 145

RESULT 22
RL10_CHLTR
ID RL10_CHLTR STANDARD; PRT; 172 AA.
AC 084319;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L10.
GN RPLJ OR RL10 OR CT317.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/Cx;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -I- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001304; AAC67910.1; -.
CC PHCT-2DPAGE; 084319; -.
DR InterPro: IPR001790; Ribosomal_L10.
DR InterPro: IPR002363; Ribosomal_L10eub.
DR Pfam: PF00466; Ribosomal_L10; 1.
DR PROSITE: PS01109; RIBOSOMAL_L10; 1.
KM Ribosomal protein: Complete proteome.
KW Ribosomal protein; Complete proteome.

SQ SEQUENCE 172 AA; 18742 MW; B67A5A3B1440667C CRC64;

Query Match 100.0%; Score 20; DB 1; Length 172;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 142 AAPM 145

RESULT 23
VEGA_BOVIN
ID VEGA_BOVIN STANDARD; PRT; 190 AA.
AC P15691;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
DE permeability factor) (VPF).
GN VEGF OR VEGFA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-47.
RX MEDLINE=90069608; PubMed=2479986;
RA Leung D.W., Cachianes G., Kuang W.-J., Goeddel D.V., Ferrara N.;
RT "Vascular endothelial growth factor is a secreted angiogenic
RT mitogen.";
RL Science 246:1306-1309(1989).
RN [2]
RP SEQUENCE OF 27-190 FROM N.A. (ISOFORMS ALPHA AND BETA).
RX MEDLINE=90121225; PubMed=2610687;
RA Tischer E., Gospodarowicz D., Mitchell R., Silva M., Schilling J.,
RA Lau K., Crisp T., Fiddes J.C., Abraham J.A.;
RT "Vascular endothelial growth factor: a new member of the platelet-
RT derived growth factor gene family.";
RL Biochem. Biophys. Res. Commun. 165:1198-1206(1989).
RN [3]
RP SEQUENCE OF 27-31.
RX MEDLINE=89286596; PubMed=2735925;
RA Ferrara N., Henzel W.J.;
RT "pituitary follicular cells secrete a novel heparin-binding growth
RT factor specific for vascular endothelial cells.";
RL Biochem. Biophys. Res. Commun. 161:851-858(1989).
CC -I- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
CC endothelial cell growth. It induces endothelial cell
CC proliferation, promotes cell migration, inhibits apoptosis, and
CC induces permeabilization of blood vessels. It binds to the
CC VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
CC heparin (By similarity).
CC -I- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
CC with PLGF (By similarity).
CC -I- SUBCELLULAR LOCATION: Secreted but remains associated to cells or
CC to the extracellular matrix unless released by heparin (By
CC similarity).
CC -I- ALTERNATIVE PRODUCTS: 2 isoforms; alpha (shown here) and beta; are
CC produced by alternative splicing.
CC -I- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M32976; AAA30502.1; -.
DR EMBL: M31836; AAA30804.1; -.

```
DR EMBL; M33750; AAA30805.1; -.
DR PIR; A33255; A33255.
DR PIR; A33787; A33787.
DR PIR; B40080; B40080.
DR HSSP; P15692; 1VGH.
DR InterPro; IPR000072; PD_growth_factor.
DR Pfam; PF00341; PDGF_1.
DR ProDom; PD001629; PD_growth_factor; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
KW Mitogen; Angiogenesis; Growth factor; Glycoprotein; Signal;
KW Heparin-binding; Alternative splicing; Multigene family.
FT SIGNAL 1 26
FT CHAIN 27 190 VASCULAR ENDOTHELIAL GROWTH FACTOR A.
FT DISULFID 51 93 BY SIMILARITY.
FT DISULFID 82 127 BY SIMILARITY.
FT DISULFID 86 129 BY SIMILARITY.
FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 139 183 MISSING (IN ISOFORM BETA).
FT VARSPLIC 184 184 R -> K (IN ISOFORM BETA).
SQ SEQUENCE 190 AA; 22310 MW; EDHF903E46E24789 CRC64;
```

```
Query Match 100.0%; Score 20; DB 1; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    |||
Db 26 AAPM 29
```

```
RESULT 24
VEGA_HORSE STANDARD; PRT; 190 AA.
ID VEGA_HORSE
AC O9GKR0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
DE permeability factor) (VPF).
GN VEGF OR VEGFA.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Miura N., Misumi K., Kawahara K., Nakashima M., Fukumitsu S.,
RA Kawabata H., Uto N., Oka T., Maruyama I., Sakamoto H.;
RT "Cloning of cDNA and high-level expression of equine vascular
RT endothelial growth factor (VEGF).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Growth factor active in angiogenesis, and endothelial
CC cell growth. Induces endothelial proliferation and vascular
CC permeability (By similarity).
CC -!- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
CC with PLGF (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted but remains associated to cells or
CC to the extracellular matrix unless released by heparin (By
CC similarity).
CC -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
```

```
DR EMBL; AB053350; BAB20890.1; -.
DR HSSP; P15692; 1VGH.
DR InterPro; IPR000072; PD_growth_factor.
DR Pfam; PF00341; PDGF; 1.
DR ProDom; PD001629; PD_growth_factor; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
KW Mitogen; Angiogenesis; Growth factor; Glycoprotein; Signal;
KW Multigene family.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 190 VASCULAR ENDOTHELIAL GROWTH FACTOR A.
FT DISULFID 51 93 BY SIMILARITY.
FT DISULFID 82 127 BY SIMILARITY.
FT DISULFID 86 129 BY SIMILARITY.
FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 190 AA; 22312 MW; 87E9E161439E5F87 CRC64;
```

```
Query Match 100.0%; Score 20; DB 1; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    |||
Db 26 AAPM 29
```

```
RESULT 25
VEGA_PIG STANDARD; PRT; 190 AA.
ID VEGA_PIG
AC P49151; O9GL52;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
DE permeability factor) (VPF).
GN VEGF OR VEGFA.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=95143284; Pubmed=7841203;
RA Sharma H.S., Tang Z.H., Gho B.C.H., Verdouw P.D.;
RT "Nucleotide sequence and expression of the porcine vascular
RT endothelial growth factor.";
RL Biochim. Biophys. Acta 1260:235-238(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Lee T., Cauty J.M.;
RT "PCR cloning of porcine cardiac vascular endothelial growth factor
RT gene.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
CC endothelial cell growth. It induces endothelial cell
CC proliferation, promotes cell migration, inhibits apoptosis, and
CC induces permeabilization of blood vessels. It binds to the
CC VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
CC heparin (By similarity).
CC -!- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
CC with PLGF (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted but remains associated to cells or
CC to the extracellular matrix unless released by heparin (By
CC similarity).
CC -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```



```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X81380; CAA57143.1; -.
DR EMBL; AF318502; AAG33064.1; -.
DR HSSP; P15692; 1VGH.
DR InterPro; IPR000072; PD_growth_factor.
DR Pfam; PF00341; PDGF; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR Mitogen; Angiogenesis; Growth factor; Glycoprotein; Signal;
KW Heparin-binding; Multigene family.
FT SIGNAL 1 26
FT CHAIN 27 190
FT DISULFID 51 93
FT DISULFID 82 127
FT DISULFID 86 129
FT DISULFID 76 76
FT DISULFID 85 85
FT CARBOHYD 100 100
FT CONFLICT 102 102
SQ SEQUENCE 190 AA; 22368 MW; 04D40B8D7913047F CRC64;

Query Match 100.0%; Score 20; DB 1; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 26 AAPM 29

RESULT 26
CBP2_ARATH STANDARD; PRT; 191 AA.
ID CBP2_ARATH
AC P30187;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 22 kDa calmodulin-like calcium-binding protein (CABP-22).
GN AT2G41090 OR TK9.14.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia; TISSUE=leaf;
RX MEDLINE=93283626; PubMed=8507824;
RA Ling V., Zielinski R.E.;
RT "Isolation of an Arabidopsis cDNA sequence encoding a 22 kDa calcium-
RT binding protein (CABP-22) related to calmodulin.";
RL Plant Mol. Biol. 22:207-214(1993).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Niernman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
```

```
CC -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN LEAVES.
CC -!- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z12136; CAA78124.1; -.
DR EMBL; AC004261; AAD12002.1; -.
DR PIR; S25125; S25125.
DR PIR; S35188; S35188.
DR HSSP; P02593; 1CDM.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand; 4.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; EFh; 4.
DR PROSITE; PS00018; EF_HAND; 4.
KW Calcium-binding; Repeat.
FT CA_BIND 21 32
FT CA_BIND 57 68
FT CA_BIND 91 102
FT CA_BIND 127 138
SQ SEQUENCE 191 AA; 21727 MW; 571F40C1E64370A8 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 191;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 175 AAPM 178

RESULT 27
ACD2_CLOPE STANDARD; PRT; 198 AA.
ID ACD2_CLOPE
AC Q8XM99;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative acyl carrier protein phosphodiesterase 2 (EC 3.1.4.14) (ACP
DE phosphodiesterase 2).
GN CPE0791.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hiraoka H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
RN [2]
RP FUNCTION: Converts holo-ACP to apo-ACP by hydrolytic cleavage of
RN the phosphopantetheine prosthetic group from ACP (By similarity).
RN CATALYTIC ACTIVITY: Holo-[acyl-carrier protein] + H(2)O = 4'-
RN phosphopantetheine + apo-[acyl-carrier protein].
RN -!- SIMILARITY: BELONGS TO THE ACP FAMILY.
RN -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```



```
CC -----
DR EMBL: AP003188; BAB80497.1; -.
KW Hypothetical protein; Hydrolase; Complete proteome.
SQ SEQUENCE 198 AA; 22383 MW; 1BFEB0B6B2EB21C3 CRC64;

Query Match          100.0%; Score 20; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
   ||||
Db 89 AAPM 92

RESULT 28
ACPD_ECOS7          STANDARD;          PRT;          200 AA.
ID ACPD_ECOS7
AC Q8X9S9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acyl carrier protein phosphodiesterase (EC 3.1.4.14) (ACP
DE phosphodiesterase).
GN ACPD OR Z2315 OR ECS2014.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / AFCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -1- FUNCTION: Converts holo-ACP to apo-ACP by hydrolytic cleavage of
CC the phosphopantetheine prosthetic group from ACP. Its
CC physiological function is not clear (By similarity).
CC -1- CATALYTIC ACTIVITY: Holo-[acyl-carrier protein] + H(2)O = 4'-
CC phosphopantetheine + apo-[acyl-carrier protein].
CC -1- SIMILARITY: BELONGS TO THE ACPD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE005365; AAG56367.1; -.
DR EMBL: AP002557; BAB35437.1; -.
KW Hydrolase; Complete proteome.
FT INIT_MET 0
SQ SEQUENCE 200 AA; 21514 MW; 5672782FF940FFCE CRC64;
Query Match          100.0%; Score 20; DB 1; Length 200;
```

```
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
   ||||
Db 92 AAPM 95

RESULT 29
ACPD_ECOLI          STANDARD;          PRT;          200 AA.
ID ACPD_ECOLI
AC P41407; P77143; Q93V21;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acyl carrier protein phosphodiesterase (EC 3.1.4.14) (ACP
DE phosphodiesterase).
GN ACPD OR B1412.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Kitakawa M., Kasai H., Baba T., Honjo A., Isono K.;
RT "Nucleotide sequence of the replication terminus region of Escherichia
RT coli.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Salto N.,
RA Sempel G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [4]
RP SEQUENCE OF 1-143 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=95206938; PubMed=7899078;
RA Moriya H., Kasai H., Isono K.;
RT "Cloning and characterization of the hrpA gene in the terC region of
RT Escherichia coli that is highly similar to the DEAH family RNA
RT helicase genes of Saccharomyces cerevisiae.";
RL Nucleic Acids Res. 23:595-598(1995).
RN [5]
RP SEQUENCE OF 1-15, AND CHARACTERIZATION.
RX MEDLINE=90368607; PubMed=2168383;
RA Fischl A.S., Kennedy E.P.;
RT "Isolation and properties of acyl carrier protein phosphodiesterase
RT of Escherichia coli.";
RL J. Bacteriol. 172:5445-5449(1990).
CC -1- FUNCTION: Converts holo-ACP to apo-ACP by hydrolytic cleavage of
CC the phosphopantetheine prosthetic group from ACP. Its
CC physiological function is not clear.
CC -1- CATALYTIC ACTIVITY: Holo-[acyl-carrier protein] + H(2)O = 4'-
CC phosphopantetheine + apo-[acyl-carrier protein].
CC
```

```
CC -i- SIMILARITY: BELONGS TO THE ACPD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D85081; BAA25408.1; -.
DR EMBL; AE000238; AAC74494.1; -.
DR EMBL; D90780; BAA15028.1; -.
DR EMBL; D90779; BAA15024.1; -.
DR EMBL; D42105; BAA07684.1; -.
DR EcoGene; EG12695; acpd.
KW Hydrolyase; Complete proteome.
FT INIT_MET 0
FT CONFLICT 22 34 DYEVEQWREKHS A -> IILLNGAKSTPR (IN
FT REF. 1 AND 4).
SQ SEQUENCE 200 AA; 21526 MW; E26A182AED432BC8 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 200;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 92 AAPM 95

RESULT 30
ACPD_SALTY
ID ACPD_SALTY STANDARD; PRT; 200 AA.
AC Q8XFP4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acyl carrier protein phosphodiesterase (EC 3.1.4.14) (ACP
DE phosphodiesterase).
GN ACPD OR STM1642 OR STY1427.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RT Nature 413:852-856(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Conerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltham T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RT Nature 413:848-852(2001).
CC -----
```

```
CC -i- FUNCTION: Converts holo-ACP to apo-ACP by hydrolytic cleavage of
CC the phosphopantetheine prosthetic group from ACP. Its
CC physiological function is not clear (By similarity).
CC -i- CATALYTIC ACTIVITY: Holo-[acyl-carrier protein] + H(2)O = 4'-
CC phosphopantetheine + apo-[acyl-carrier protein].
CC -i- SIMILARITY: BELONGS TO THE ACPD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008772; AAL20560.1; -.
DR EMBL; AL627270; CAD01689.1; -.
DR StyGene; SG?27270; acpd.
KW Hydrolyase; Complete proteome.
FT INIT_MET 0
FT CONFLICT 0 BY SIMILARITY.
SQ SEQUENCE 200 AA; 21496 MW; 8EC14DA11C080062 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 200;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 92 AAPM 95

RESULT 31
ACPD_YERPE
ID ACPD_YERPE STANDARD; PRT; 200 AA.
AC Q8ZE60;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acyl carrier protein phosphodiesterase (EC 3.1.4.14) (ACP
DE phosphodiesterase).
GN ACPD OR YPO2323.
OS Yersinia pestis.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltham T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RA "Genome sequence of Yersinia pestis, the causative agent of plague.";
RT Nature 413:523-527(2001).
RN [2]
RP FUNCTION: Converts holo-ACP to apo-ACP by hydrolytic cleavage of
CC the phosphopantetheine prosthetic group from ACP. Its
CC physiological function is not clear (By similarity).
CC -i- CATALYTIC ACTIVITY: Holo-[acyl-carrier protein] + H(2)O = 4'-
CC phosphopantetheine + apo-[acyl-carrier protein].
CC -i- SIMILARITY: BELONGS TO THE ACPD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```

DR EMBL; AJ414152; CAC91128.1; -.
KW Hydrolase; Complete proteome.
FT INIT_MET 0 BY SIMILARITY.
SQ SEQUENCE 200 AA; 21468 MW; 0512BF52800808F CRC64;

Query Match 100.0%; Score 20; DB 1; Length 200;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
   1111
Db 92 AAPM 95

RESULT 32
ACD3_PSEAE STANDARD; PRT; 213 AA.
ID ACD3_PSEAE
AC Q9H217;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative acyl carrier protein phosphodiesterase 3 (EC 3.1.4.14) (ACP
DE phosphodiesterase 3).
GN PA3223.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -1- FUNCTION: Converts holo-ACP to apo-ACP by hydrolytic cleavage of
CC the phosphopantetheine prosthetic group from ACP (By similarity).
CC -1- CATALYTIC ACTIVITY: Holo-(acyl-carrier protein) + H(2)O = 4'-
CC phosphopantetheine + apo-[acyl-carrier protein].
CC -1- SIMILARITY: BELONGS TO THE ACPD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE004745; AAG06611.1; -.
KW Hypothetical protein; Hydrolase; Complete proteome.
SQ SEQUENCE 213 AA; 23898 MW; DBE4B84DFD627E7B CRC64;

Query Match 100.0%; Score 20; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
   1111
Db 93 AAPM 96

RESULT 33
VEGA_CANFA STANDARD; PRT; 214 AA.
ID VEGA_CANFA
AC Q9MYV3; Q9XSF5; Q9XSF4; Q9XSF3;
DT 15-JUN-2002 (Rel. 41, Created)

```

```

DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
DE permeability factor) (VPF).
GN VEGF OR VEGFA.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM VEGF188).
RX MEDLINE=20125516; PubMed=10661874;
RA Scheidegger P., Weiglhofer W., Suarez S., Kaser-Holtz B., Steiner R.,
RA Ballmer-Hofer K., Jaussi R.;
RT "Vascular endothelial growth factor (VEGF) and its receptors in tumor-
RT bearing dogs.";
RL Biol. Chem. 380:1449-1454(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS VEGF188; VEGF-182 AND VEGF-164).
RC TISSUE=Heart;
RA Jingjing L., Roque R.S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and
CC endothelial cell growth. It induces endothelial cell
CC proliferation, promotes cell migration, inhibits apoptosis, and
CC induces permeabilization of blood vessels. It binds to the
CC VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and
CC heparin (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer
CC with PlGF (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted but remains associated to cells or
CC to the extracellular matrix unless released by heparin (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS: At least 3 isoforms; VEGF188 (shown here),
CC VEGF182 and VEGF164; are produced by alternative splicing.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ133758; CAB82426.1; -.
DR EMBL; AF133250; AAD29684.1; -.
DR EMBL; AF133249; AAD29683.1; -.
DR EMBL; AF133248; AAD29682.1; -.
DR HSSP; P15692; 1VGH.
DR InterPro; IPR000072; PD_growth_factor.
DR Pfam; PF00341; PDGF; 1.
DR Prodom; PD001629; PD_growth_factor; 1.
DR SMART; SMD0141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
KW Mitogen; Angiogenesis; Growth factor; Glycoprotein; Signal;
KW Heparin-binding; Alternative splicing; Multigene family.
FT SIGNAL 1 26
FT CHAIN 27 214 VASCULAR ENDOTHELIAL GROWTH FACTOR A.
FT DISULFID 51 93 BY SIMILARITY.
FT DISULFID 82 127 BY SIMILARITY.
FT DISULFID 86 129 BY SIMILARITY.
FT DISULFID 76 76 INTERCHAIN (BY SIMILARITY).
FT DISULFID 85 85 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 100 100 K -> N (IN ISOFORM VEGF-164).
FT VARSPLIC 140 140 MISSING (IN ISOFORM VEGF-164).
FT VARSPLIC 141 164 MISSING (IN ISOFORM VEGF-164).
FT VARSPLIC 159 164 MISSING (IN ISOFORM VEGF-182).
FT CONFLICT 143 143 I -> V (IN REF. 2).
FT CONFLICT 161 161 P -> S (IN REF. 2).
SQ SEQUENCE 214 AA; 25175 MW; 0AC980A158CA4B27 CRC64;

```

```
Query Match          100.0%; Score 20; DB 1; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPM 4
      ||||
Db      26 AAPM 29

RESULT 34
RS5_DROME STANDARD; PRT; 228 AA.
ID RS5_DROME
AC 024186; Q9VX78;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 40S ribosomal protein S5.
DE RP55 OR M(1)5D OR CG8922.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97032939; PubMed=8878687;
RA McKim K.S., Dahmus J., Hawley R.S.;
RT "Cloning of the Drosophila melanogaster meiotic recombination gene
RL mei-218: a genetic and molecular analysis of interval 15E.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenlison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Spier E., Spradling A.C., Stapleton M., Skupski M.P., Smith T.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; U48394; AAB61633.1; -.
DR EMBL; AE003504; AAF48700.1; -.
DR HSSP; P22744; 1HUS.
DR FlyBase; FBgn0002590; Rp55.
DR InterPro; IPR000235; Ribosomal_S7.
DR Pfam; PF00177; Ribosomal_S7; 1.
DR ProDom; PD000817; Ribosomal_S7; 1.
DR TIGRFAMs; TIGR01028; S7_S5_E_A; 1.
DR PROSITE; PS00052; RIBOSOMAL_S7; 1.
KW Ribosomal protein.
SQ SEQUENCE 228 AA; 25434 MW; 9206E1D8FE470C4C CRC64;
```

```
Query Match          100.0%; Score 20; DB 1; Length 228;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPM 4
      ||||
Db      16 AAPM 19
```

```
RESULT 35
VMAT_VSVSJ
ID VMAT_VSVSJ STANDARD; PRT; 229 AA.
AC P03519;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Matrix protein.
GN M.
OS Vesicular stomatitis virus (strain San Juan).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Vesiculovirus.
OX NCBI_TaxID=11285;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82010868; PubMed=6268840;
RA Rose J.K., Gallione C.J.;
RT "Nucleotide sequences of the mRNA's encoding the vesicular stomatitis
RT virus G and M proteins determined from cDNA clones containing the
RT complete coding regions.";
RL J. Virol. 39:519-528(1981).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; J02428; AAA48369.1; -.
DR PIR; A04111; MFVN.
KW Matrix protein; Envelope protein.
SQ SEQUENCE 229 AA; 26094 MW; C89E2ACA4365B847 CRC64;
```

```
Query Match          100.0%; Score 20; DB 1; Length 229;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPM 4
      ||||
Db      185 AAPM 188
```



```
RESULT 36
YXHK_CAEEL          STANDARD;          PRT;          229 AA.
ID YXHK_CAEEL
AC Q20728;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein F53F4.3 in chromosome V.
GN F53F4.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Barlow K.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 CAP-GLY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 277663; CAB01212.1; -
DR WormPep; F53F4.3; CEI0958.
DR InterPro: IPR000938; CAP-GLY.
DR Pfam; PF01302; CAP_GLY; 1.
DR PROSITE; PS00845; CAP_GLY_1; 1.
DR PROSITE; PS50245; CAP_GLY_2; 1.
KW Hypothetical protein; Cytoskeleton; Microtubules.
FT DOMAIN 170 212 CAP-GLY.
SQ SEQUENCE 229 AA; 25440 MW; C465365DAE378A0F CRC64;

Query Match          100.0%; Score 20; DB 1; Length 229;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 128 AAPM 131

RESULT 37
YR01_CAEEL
ID YR01_CAEEL          STANDARD;          PRT;          231 AA.
AC Q10014;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein T25E4.1 in chromosome II precursor.
GN T25E4.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Johnson D.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Waterston R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: STRONG, TO C.ELEGANS D2096.6.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U23411; AAC46730.2; -
DR WormPep; T25E4.1; CE28493.
KW Hypothetical protein; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 231 HYPOTHETICAL PROTEIN T25E4.1.
SQ SEQUENCE 231 AA; 24613 MW; 8B32AF97C3F7BB08 CRC64;

Query Match          100.0%; Score 20; DB 1; Length 231;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 84 AAPM 87

RESULT 38
VEGA_HUMAN
ID VEGA_HUMAN          STANDARD;          PRT;          232 AA.
AC P15692; Q16889; O60720; O75875; Q9UL23; Q9UH58; Q9HIW9; Q9HIW8;
AC Q96NW5;
DT 01-APR-1990 (Rel. 14, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor A precursor (VEGF-A) (Vascular
DE permeability factor) (VPF).
GN VEGF OR VEGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS VEGF189 AND VEGF165).
RX MEDLINE=90069608; PubMed=2479986;
RA Leung D.W., Cachianes G., Kuang W.-J., Goeddel D.V., Ferrara N.;
RT "Vascular endothelial growth factor is a secreted angiogenic
RT mitogen.";
RL Science 246:1306-1309(1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM VEGF189), AND PARTIAL SEQUENCE.
RX MEDLINE=90069609; PubMed=2479987;
RA Keck P.J., Hauser S.D., Krivi G., Sanzo K., Warren T., Feder J.,
RA Connolly D.T.;
RT "Vascular permeability factor, an endothelial cell mitogen related to
RT PDGF.";
RL Science 246:1309-1312(1989).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM VEGF189).
RX MEDLINE=91268072; PubMed=1711045;
RA Fischer E., Mitchell R., Hartman T., Silva M., Gospodarowicz D.,
RA Fiddes J.C., Abraham J.A.;
RT "The human gene for vascular endothelial growth factor. Multiple
RT protein forms are encoded through alternative exon splicing.";
RL J. Biol. Chem. 266:11947-11954(1991).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM VEGF206).
RX MEDLINE=92168017; PubMed=1791831;
RA Houck K.A., Ferrara N., Winer J., Cachianes G., Li B., Leung D.W.;
RT "The vascular endothelial growth factor family: Identification of a
RT fourth molecular species and characterization of alternative splicing
RT of RNA.";
RL Mol. Endocrinol. 5:1806-1814(1991).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM VEGF165).
RX MEDLINE=92231879; PubMed=1567395;
```


| | | | |
|------|---|----|--|
| RA | Weindel K., Marne D., Welch H.A.; | RT | "Human vascular permeability factor. Isolation from U937 cells."; |
| RT | "AIDS-associated Kaposi's sarcoma cells in culture express vascular | RL | J. Biol. Chem. 264:20017-20024(1989). |
| RT | endothelial growth factor."; | RN | [16] |
| RL | Biochem. Biophys. Res. Commun. 183:1167-1174(1992). | RP | SEQUENCE OF 27-41. |
| RN | [6] | RX | MEDLINE-93145946; Pubmed=7678805; |
| RP | SEQUENCE FROM N.A. (ISOFORM VEGF145). | RA | Fiebig B.L., Jaeger B., Schoellmann C., Weindel K., Wilting J., |
| RX | MEDLINE-97207275; Pubmed=9054410; | RA | Kochs G., Marne D., Hug H., Welch H.A.; |
| RA | Poltorak Z., Cohen T., Sivan R., Kandelis Y., Spira G., Vlodaysky I., | RT | "Synthesis and assembly of functionally active human vascular |
| RA | Keshet E., Neufeld G.; | RT | endothelial growth factor homodimers in insect cells."; |
| RT | "VEGF145, a secreted vascular endothelial growth factor isoform that | RL | Eur. J. Biochem. 211:19-26(1993). |
| RT | binds to extracellular matrix."; | RN | [17] |
| RL | J. Biol. Chem. 272:7151-7158(1997). | RP | X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 34-135. |
| RN | [7] | RX | MEDLINE-97352774; Pubmed=9207067; |
| RP | SEQUENCE FROM N.A. (ISOFORM VEGF183). | RA | Muller Y.A., Li B., Christinger H.W., Wells J.A., Cunningham B.C., |
| RC | TISSUE=Kidney; | RA | de Vos A.M.; |
| RX | MEDLINE-99096474; Pubmed=9878851; | RT | "Vascular endothelial growth factor: crystal structure and functional |
| RA | Lei J., Jiang A., Pei D.; | RT | mapping of the kinase domain receptor binding site."; |
| RT | "Identification and characterization of a new splicing variant of | RL | Proc. Natl. Acad. Sci. U.S.A. 94:7192-7197(1997). |
| RT | vascular endothelial growth factor: VEGF183."; | RN | [18] |
| RL | Biochim. Biophys. Acta 1443:400-406(1998). | RP | X-RAY CRYSTALLOGRAPHY (1.93 ANGSTROMS) OF 34-135. |
| RN | [8] | RX | MEDLINE-98035455; Pubmed=9351807; |
| RP | SEQUENCE FROM N.A. (ISOFORM VEGF165). | RA | Muller Y.A., Christinger H.W., Keyt B.A., de Vos A.M.; |
| RC | TISSUE-Breast; | RT | "The crystal structure of vascular endothelial growth factor (VEGF) |
| RX | MEDLINE-98119755; Pubmed=9450968; | RT | refined to 1.93-A resolution: multiple copy flexibility and receptor |
| RA | Ciafey K.P., Shih S.-C., Mullen A., Dziennis S., Cusick J.L., | RT | binding."; |
| RA | Abrams K.R., Lee S.W., Delmar M.; | RL | Structure 5:1325-1338(1997). |
| RT | "Identification of a human VPE/VEGF 3' untranslated region mediating | RN | [19] |
| RT | hypoxia-induced mRNA stability."; | RP | X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 39-134. |
| RL | Mol. Biol. Cell 9:469-481(1998). | RX | MEDLINE-99119204; Pubmed=9922142; |
| RN | [9] | RA | Wiesmann C., Christinger H.W., Cochran A.G., Cunningham B.C., |
| RP | SEQUENCE OF 114-209 FROM N.A. (ISOFORM VEGF183). | RA | Fairbrother W.J., Keenan C.J., Meng G., de Vos A.M.; |
| RC | TISSUE-Retina; | RT | "Crystal structure of the complex between VEGF and a receptor-blocking |
| RX | MEDLINE-99165303; Pubmed=10067980; | RT | peptide."; |
| RA | Jingjing L., Xue Y., Agarwal N., Roque R.S.; | RL | Biochemistry 37:17765-17772(1998). |
| RT | "Human Muller cells express VEGF183, a novel spliced variant of | RN | [20] |
| RT | vascular endothelial growth factor."; | RP | STRUCTURE BY NMR OF 34-135. |
| RL | Invest. Ophthalmol. Vis. Sci. 40:752-759(1999). | RX | MEDLINE-97477915; Pubmed=9336848; |
| RP | [10] | RA | Fairbrother W.J., Champe M.A., Christinger H.W., Keyt B.A., |
| RN | SEQUENCE FROM N.A. (ISOFORM VEGF165). | RA | Starovasnik M.A.; |
| RC | TISSUE-Hemangioidendelioma; | RT | "1H, 13C, and 15N backbone assignment and secondary structure of the |
| RA | Murata H., Fukushima J., Hattori S., Okuda K., Yanagi H.; | RT | receptor-binding domain of vascular endothelial growth factor."; |
| RT | "Human cDNA for the vascular endothelial growth factor isoform | RL | protein Sci. 6:2250-2260(1997). |
| RT | VEGF165."; | RN | [21] |
| RL | Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases. | RP | STRUCTURE BY NMR OF 137-215. |
| RP | [11] | RX | MEDLINE-98298440; Pubmed=9634701; |
| RN | SEQUENCE FROM N.A. (ISOFORM VEGF148). | RA | Fairbrother W.J., Champe M.A., Christinger H.W., Keyt B.A., |
| RC | TISSUE-Renal glomerulus; | RA | Starovasnik M.A.; |
| RX | MEDLINE-99394945; Pubmed=10464055; | RT | "Solution structure of the heparin-binding domain of vascular |
| RA | Whittle C.J., Gillespie K.M., Harrison R., Mathieson P.W., | RT | endothelial growth factor."; |
| RA | Harper S.J.; | RL | Structure 6:637-648(1998). |
| RT | "Heterogeneous vascular endothelial growth factor (VEGF) isoform mRNA | RN | [22] |
| RT | and receptor mRNA expression in human glomeruli, and the | RP | FUNCTION. |
| RT | identification of VEGF148 mRNA, a novel truncated splice variant."; | RX | MEDLINE-21320570; Pubmed=11427521; |
| RL | Clin. Sci. 97:303-312(1999). | RA | Murphy J.F., Fitzgerald D.J.; |
| RN | [12] | RT | "Vascular endothelial growth factor induces cyclooxygenase-dependent |
| RP | SEQUENCE FROM N.A. (ISOFORM VEGF121). | RL | proliferation of endothelial cells via the VEGF-2 receptor."; |
| RA | Sato J.D., Whitney R.G.; | CC | FASEB J. 15:1667-1669(2001). |
| RT | "Human cDNA for vascular endothelial growth factor isoform VEGF121."; | CC | -I- FUNCTION: Growth factor active in angiogenesis, vasculogenesis and |
| RT | Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases. | CC | endothelial cell growth. It induces endothelial cell |
| RN | [13] | CC | proliferation, promotes cell migration, inhibits apoptosis, and |
| RP | SEQUENCE FROM N.A. | CC | induces permeabilization of blood vessels. It binds to the |
| RP | Williams S.; | CC | VEGFR1/Flt-1 and VEGFR2/Kdr receptors and to heparan sulfate and |
| RL | Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases. | CC | heparin. Neupilin-1 binds isoforms VEGF-165 and VEGF-145. |
| RN | [14] | CC | -I- SUBUNIT: Homodimer; disulfide-linked. Also found as heterodimer |
| RP | SEQUENCE OF 23-232 FROM N.A. (VEGF165). | CC | with PlGF (by similarity). |
| RA | Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L., | CC | -I- SUBCELLULAR LOCATION: VEGF121 is acidic and freely secreted. |
| RA | Poel C.L., Toth E.J., Yi Q., Nickerson D.A.; | CC | VEGF165 is more basic, has heparin-binding properties and, |
| RP | Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases. | CC | although a significant proportion remains cell-associated, most is |
| RP | [15] | CC | freely secreted. VEGF189 is very basic; it is cell-associated |
| RN | PRELIMINARY SEQUENCE OF 27-36; 43-50 AND 59-81. | CC | after secretion and is bound avidly by heparin and the |
| RX | MEDLINE-90062112; Pubmed=2584205; | CC | extracellular matrix, although it may be released as a soluble |
| RA | Connolly D.T., Olander J.V., Heuvelman D., Nelson R., Monsell R., | CC | form by heparin, heparinase or plasmin. |
| RA</ | | | |

```
CC VEGF183, VEGF165/VEGF, VEGF148, VEGF145 and VEGF121; may be
CC produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: The VEGF189, VEGF-165 and VEGF-121 isoforms
CC are widely expressed, whereas the VEGF206 and VEGF-145 are
CC uncommon.
CC -!- INDUCTION: Regulated by growth factors, cytokines, gonadotropins,
CC nitric oxide, hypoxia, hypoglycemia and oncogenic mutations.

Query Match          100.0%; Score 20; DB 1; Length 232;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
    ||||
Db 26 AAPM 29

RESULT 39
YY26_MYCTU
ID YY26_MYCTU STANDARD; PRT; 232 AA.
AC Q50702;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical PPE-family protein RV3426.
GN RV3426 OR MTCY78.03C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Ruster S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RT Nature 393:537-544(1998).
RL Nature 393:537-544(1998).
-!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 277165; CAB01030.1; -.
DR Tuberculist; RV3426; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 232 AA; 25872 MW; D76512D49EB272C6 CRC64;
```

```
ID VMAT_VSVIG STANDARD; PRT; 237 AA.
AC P04876;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Matrix protein.
GN M.
OS Vesicular stomatitis virus (serotype Indiana / strain Glasgow).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Vesiculovirus.
OX NCBI_TaxID=11278;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86142630; PubMed=3005478;
RA Vandepol S.B., Holland J.J.;
RT "Evolution of vesicular stomatitis virus in athymic nude mice:
RT mutations associated with natural killer cell selection.";
RL J. Gen. Virol. 67:441-451(1986).
RN [2]
RP SEQUENCE OF 1-229 FROM N.A.
RX MEDLINE=86062892; PubMed=2999421;
RA Gopalakrishna Y., Lenard J.;
RT "Sequence alterations in temperature-sensitive M-protein mutants
RT (complementation group III) of vesicular stomatitis virus.";
RL J. Virol. 56:655-659(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X04452; CAA28051.1; -.
DR EMBL; M11754; AAA48444.1; -.
DR PIR; A04112; MFVNGG.
KW Matrix protein; Envelope protein.
FT CONFLICT 32 32 S -> N (IN REF. 2).
FT CONFLICT 64 64 S -> F (IN REF. 2).
FT CONFLICT 215 215 E -> K (IN REF. 2).
FT CONFLICT 228 228 S -> F (IN REF. 2).
SQ SEQUENCE 237 AA; 26824 MW; C1DF0CF71510959D CRC64;
```

```
Query Match          100.0%; Score 20; DB 1; Length 237;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
    ||||
Db 185 AAPM 188

Search completed: December 6, 2002, 13:28:33
Job time : 5.55556 secs
```

RESULT 40
VMAT_VSVIG

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:25:09 ; Search time 19.5556 Seconds
(without alignments)
42.146 Million cell updates/sec

Title: US-10-033-526-2
Perfect-score: 20
Sequence: 1 AAPM 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 20 | 100.0 | 23 | 10 Q94AM6 | Q94am6 arabidopsis |
| 2 | 20 | 100.0 | 30 | 4 Q96TC2 | Q96tc2 homo sapien |
| 3 | 20 | 100.0 | 47 | 4 Q969N8 | Q969n8 homo sapien |
| 4 | 20 | 100.0 | 53 | 12 Q36586 | Q36586 hepatitis e |
| 5 | 20 | 100.0 | 53 | 12 Q36589 | Q36589 hepatitis e |
| 6 | 20 | 100.0 | 53 | 12 Q36592 | Q36592 hepatitis e |
| 7 | 20 | 100.0 | 53 | 12 Q36583 | Q36583 hepatitis e |
| 8 | 20 | 100.0 | 53 | 12 Q36598 | Q36598 hepatitis e |
| 9 | 20 | 100.0 | 59 | 12 Q9W7U2 | Q9w7u2 hepatitis e |
| 10 | 20 | 100.0 | 61 | 6 Q9TS81 | Q9ts81 camelus dro |
| 11 | 20 | 100.0 | 71 | 10 Q8S6E6 | Q8s6e6 oryza sativ |
| 12 | 20 | 100.0 | 76 | 6 Q02709 | Q02709 monodelphis |
| 13 | 20 | 100.0 | 76 | 8 Q03382 | Q03382 salmo salar |
| 14 | 20 | 100.0 | 76 | 11 Q88946 | Q88946 mus musculu |
| 15 | 20 | 100.0 | 77 | 10 Q8S7E4 | Q8s7e4 oryza sativ |
| 16 | 20 | 100.0 | 79 | 5 Q94188 | Q94188 caenorhabdi |

| | | | | | |
|----|----|-------|----|-----------|---------------------|
| 17 | 20 | 100.0 | 83 | 12 Q91ML0 | Q91ml0 reovirus sp |
| 18 | 20 | 100.0 | 83 | 12 Q91MK9 | Q91mk9 reovirus sp |
| 19 | 20 | 100.0 | 83 | 12 Q91MK8 | Q91mk8 reovirus sp |
| 20 | 20 | 100.0 | 83 | 12 Q91MK7 | Q91mk7 reovirus sp |
| 21 | 20 | 100.0 | 83 | 12 Q91MK6 | Q91mk6 reovirus sp |
| 22 | 20 | 100.0 | 83 | 16 Q54062 | Q54062 rhizobium m |
| 23 | 20 | 100.0 | 84 | 2 Q86845 | Q86845 streptomyce |
| 24 | 20 | 100.0 | 88 | 16 Q8U5M7 | Q8u5m7 agrobacteri |
| 25 | 20 | 100.0 | 89 | 12 Q9DUJ1 | Q9dja1 foot-and-mo |
| 26 | 20 | 100.0 | 92 | 3 Q96UH7 | Q96uh7 paracoccidi |
| 27 | 20 | 100.0 | 94 | 16 Q9RUE9 | Q9rje9 streptomyce |
| 28 | 20 | 100.0 | 98 | 8 Q47762 | Q47762 myxocyprius |
| 29 | 20 | 100.0 | 98 | 8 Q47775 | Q47775 xenocyprius |
| 30 | 20 | 100.0 | 98 | 8 Q47781 | Q47781 xenocyprius |
| 31 | 20 | 100.0 | 98 | 8 Q9G6S9 | Q9g6s9 sardinops m |
| 32 | 20 | 100.0 | 98 | 8 Q9T919 | Q9t919 aristichthy |
| 33 | 20 | 100.0 | 98 | 8 Q9T911 | Q9t911 megalobrama |
| 34 | 20 | 100.0 | 98 | 8 Q9MIA1 | Q9mia1 typhlonecte |
| 35 | 20 | 100.0 | 98 | 8 Q9G3Z2 | Q9g3z2 coltus bair |
| 36 | 20 | 100.0 | 98 | 8 Q9G3Y9 | Q9g3y9 coltus caro |
| 37 | 20 | 100.0 | 98 | 8 Q9G3Y6 | Q9g3y6 coltus hyps |
| 38 | 20 | 100.0 | 98 | 8 Q9G3Y3 | Q9g3y3 coltus cogn |
| 39 | 20 | 100.0 | 98 | 8 Q94YP8 | Q94yp8 osteoglossu |
| 40 | 20 | 100.0 | 98 | 8 Q94TB1 | Q94tb1 neoscopelus |
| 41 | 20 | 100.0 | 98 | 8 Q94TP98 | Q94tp98 myctophom a |
| 42 | 20 | 100.0 | 98 | 8 Q94T85 | Q94t85 diaphus spl |
| 43 | 20 | 100.0 | 98 | 8 Q94SCL | Q94scl hellicolenus |
| 44 | 20 | 100.0 | 98 | 8 Q48322 | Q48322 xenocyprius |
| 45 | 20 | 100.0 | 98 | 8 Q48366 | Q48366 xenocyprius |

ALIGNMENTS

RESULT 1
ID Q94AM6 PRELIMINARY: PRT: 23 AA.
AC Q94AM6:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 2.8 kDa protein.
GN T6L1.23.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koeseema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shin P., Southwick A., Tracy S.E., Shinzaki K., Davis R.W.,
RA Ecker J.R., Theologis A.,
RT "Full length cDNA of gene T6L1.23 (not previously annotated).";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY045928; AAK76602.1; -
KW Hypothetical protein.
SQ SEQUENCE 23 AA; 2756 MW; DFL4E30FC0FBF9BC CRC64;

Query Match 100.0%; Score 20; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAPM 4
Db 15 AAPM 18

RESULT 2

Q96TC2 PRELIMINARY; PRT; 30 AA.
AC Q96TC2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Brain-derived neurotrophic factor (Fragment).
GN BDNF 2-5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Aoyama M., Asai K., Kato T., Miura Y., Kokubo M., Kawai Y.,
RA Miyachi T., Yokoi T., Tsuiki H.;
RT "partial sequence of human BDNF mRNA."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038671; BAB55546.1; -.
FT NON_TER 30
SQ SEQUENCE 30 AA; 3323 MW; B33237A993FF4BD0 CRC64;

Query Match 100.0%; Score 20; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 18 AAPM 21

RESULT 3

Q969N8 PRELIMINARY; PRT; 47 AA.
AC Q969N8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Brain-derived neurotrophic factor (Fragment).
GN BDNF 1-5 OR BDNF 3-5 OR BDNF 4-5 OR BDNF 41-5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Aoyama M., Asai K., Kato T., Miura Y., Kokubo M., Kawai Y.,
RA Miyachi T., Yokoi T., Tsuiki H.;
RT "partial sequence of human BDNF mRNA."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Aoyama M., Asai K., Kato T., Miura Y., Kokubo M., Kawai Y.,
RA Miyachi T., Yokoi T., Tsuiki H.;
RT "partial sequence of human BDNF mRNA."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038670; BAB55545.1; -.
DR EMBL; AB038672; BAB55547.1; -.
DR EMBL; AB038673; BAB55548.1; -.
DR EMBL; AB038674; BAB55549.1; -.
FT NON_TER 47
SQ SEQUENCE 47 AA; 5076 MW; A9D3DC22639759B1 CRC64;

Query Match 100.0%; Score 20; DB 4; Length 47;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 18 AAPM 21

RESULT 4

O36586 PRELIMINARY; PRT; 53 AA.
AC O36586;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORF-3 (Fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOROCCO/F12;
RX MEDLINE=97475412; PubMed=9334924;
RA Chatterjee R., Tsarev S., Pillot J., Coursaget P., Emerson S.U.,
RA Purcell R.H.;
RT "African strains of hepatitis E virus that are distinct from Asian strains."
RL J. Med. Virol. 53:139-144(1997).
DR EMBL; AF010418; AAB66533.1; -.
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2; 1.
FT NON_TER 53
SQ SEQUENCE 53 AA; 5405 MW; 638F95EEAF7E62D CRC64;

Query Match 100.0%; Score 20; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 7 AAPM 10

RESULT 5

O36589 PRELIMINARY; PRT; 53 AA.
AC O36589;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORF-3 (Fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOROCCO/F13;
RX MEDLINE=97475412; PubMed=9334924;
RA Chatterjee R., Tsarev S., Pillot J., Coursaget P., Emerson S.U.,
RA Purcell R.H.;
RT "African strains of hepatitis E virus that are distinct from Asian strains."
RL J. Med. Virol. 53:139-144(1997).
DR EMBL; AF010419; AAB66536.1; -.
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2; 1.
FT NON_TER 53
SQ SEQUENCE 53 AA; 5405 MW; 638F95EEAF7E62D CRC64;

Query Match 100.0%; Score 20; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 7 AAPM 10

RESULT 6

```
036592
ID 036592      PRELIMINARY:      PRT:      53 AA.
AC 036592:
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORF-3 (Fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOROCCO/E23;
RX MEDLINE=97475412; PubMed=9334924;
RA Chatterjee R., Tsarev S., Pillot J., Coursaget P., Emerson S.U.,
RA Purcell R.H.;
RT "African strains of hepatitis E virus that are distinct from Asian
RT strains.";
RL J. Med. Virol. 53:139-144(1997).
DR EMBL; AF010420; AAB66539.1; -.
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2; 1.
FT NON_TER      53
SQ SEQUENCE      53 AA; 5405 MW; 638F95EAF7E62D CRC64;
```

```
Query Match      100.0%; Score 20; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    ||||
Db 7 AAPM 10
```

```
RESULT 7
036583      PRELIMINARY:      PRT:      53 AA.
ID 036583:
AC 036583:
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORF-3 (Fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIRGIZIA;
RX MEDLINE=97475412; PubMed=9334924;
RA Chatterjee R., Tsarev S., Pillot J., Coursaget P., Emerson S.U.,
RA Purcell R.H.;
RT "African strains of hepatitis E virus that are distinct from Asian
RT strains.";
RL J. Med. Virol. 53:139-144(1997).
DR EMBL; AF010417; AAB66530.1; -.
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2; 1.
FT NON_TER      53
SQ SEQUENCE      53 AA; 5405 MW; 638F95EAF7E62D CRC64;
```

```
Query Match      100.0%; Score 20; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    ||||
Db 7 AAPM 10
```

RESULT 8
036598

```
ID 036598      PRELIMINARY:      PRT:      53 AA.
AC 036598:
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORF-3 (Fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UZBEKISTAN;
RX MEDLINE=97475412; PubMed=9334924;
RA Chatterjee R., Tsarev S., Pillot J., Coursaget P., Emerson S.U.,
RA Purcell R.H.;
RT "African strains of hepatitis E virus that are distinct from Asian
RT strains.";
RL J. Med. Virol. 53:139-144(1997).
DR EMBL; AF010422; AAB66545.1; -.
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2; 1.
FT NON_TER      53
SQ SEQUENCE      53 AA; 5405 MW; 638F95EAF7E62D CRC64;
```

```
Query Match      100.0%; Score 20; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    ||||
Db 7 AAPM 10
```

```
RESULT 9
09W702      PRELIMINARY:      PRT:      59 AA.
ID 09W702:
AC 09W702:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE ORF-3 (Fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M7/ALGERIA, AND T3;
RX MEDLINE=98071272; PubMed=9407381;
RA van Cuyck-Gandre H., Zhang H.Y., Tsarev S.A., Clements N.J.,
RA Cohen S.J., Caudill J.D., Buisson Y., Coursaget P., Warren R.L.,
RA Longer C.F.;
RT "Characterization of hepatitis E virus (HEV) from Algeria and Chad by
RT partial genome sequence.";
RL J. Med. Virol. 53:340-347(1997).
DR EMBL; AF001276; AAB94812.1; -.
DR EMBL; AF001275; AAB94809.1; -.
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2; 1.
FT NON_TER      59
SQ SEQUENCE      59 AA; 5906 MW; AC9B75C5F23E338F CRC64;
```

```
Query Match      100.0%; Score 20; DB 12; Length 59;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPM 4
    ||||
Db 7 AAPM 10
```

RESULT 10

Q9TS81 ID Q9TS81 PRELIMINARY; PRT; 61 AA.
AC Q9TS81;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Glycerol ester HYDROLAS (EC 3.1.1.3) (Fragments).
OS Camelus dromedarius (Dromedary) (Arabian camel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX NCBI_TaxID=9838;
RN [1]
RP SEQUENCE.
RX MEDLINE=94297034; PubMed=8025121;
RA Mejdoub H., Reinbolt J., Gargouri Y.;
RT "Dromedary pancreatic lipase: purification and structural
RT properties.";
RL Biochim. Biophys. Acta 1213:119-126(1994).
DR HSSP; P29183; IHP.L.
FT NON_TER 1 1
FT NON_CONS 20 21
FT NON_CONS 46 47
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6593 MW; DFA1874550BAD5F2 CRC64;

Query Match 100.0%; Score 20; DB 6; Length 61;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 29 AAPM 32

RESULT 11
Q8S6E6 ID Q8S6E6 PRELIMINARY; PRT; 71 AA.
AC Q8S6E6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 7.3 kDa protein.
GN OSJNB0024B16.9.
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan O., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tsitrin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pai G.,
RA VanAken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNB0024B16 genomic sequence.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC093017; AAL83624.1; -.
KW Hypothetical protein.
SQ SEQUENCE 71 AA; 7299 MW; F5DCD01F81C7D669 CRC64;

Query Match 100.0%; Score 20; DB 10; Length 71;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 16 AAPM 19

RESULT 12
002709

ID 002709 PRELIMINARY; PRT; 76 AA.
AC 002709;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Brain-derived neurotrophic factor (Fragment).
GN BDNF.
OS Monodelphis domestica (Short-tailed grey opossum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Monodelphis.
OX NCBI_TaxID=13616;
RN [1]
RP SEQUENCE FROM N.A.
RA Vischer H.A.;
RT "BDNF is expressed at the crush-site after spinal cord-lesion in
RT neonatal opossum (Monodelphis domestica).";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U95024; AAB53633.1; -.
FT NON_TER 1 1
FT NON_TER 76 76
SQ SEQUENCE 76 AA; 8421 MW; 37C9EDEDBF076259 CRC64;

Query Match 100.0%; Score 20; DB 6; Length 76;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 9 AAPM 12

RESULT 13
O03382 ID O03382 PRELIMINARY; PRT; 76 AA.
AC O03382;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE NADH dehydrogenase subunit 4L (Fragment).
OS Salmo salar (Atlantic salmon).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SHIBETSU SALMON PARK;
RX MEDLINE=97159563; PubMed=9007022;
RA Oohara I., Sawano K., Okazaki T.;
RT "Mitochondrial DNA sequence analysts of the masu salmon-phylogeny in
RT the genus Oncorhynchus.";
RL Mol. Phylogenet. Evol. 7:71-78(1997).
DR EMBL; D84148; BAA20160.1; -.
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; Oxidored_g2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion.
FT NON_TER 76 76
SQ SEQUENCE 76 AA; 8093 MW; E0A2B08AFE297F56 CRC64;

Query Match 100.0%; Score 20; DB 8; Length 76;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 58 AAPM 61

RESULT 14
088946

ID 088946 PRELIMINARY; PRT; 76 AA.
AC 088946;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, last annotation update)
DE Antisense product of high-affinity glutamate transporter EAAC1 and EAAC2.
DE Antisense product of high-affinity glutamate transporter EAAC1 and EAAC2.
GN ASECL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=BRAIN;
RA Jin X., Peng J., Fei J., Guo L.;
RT "A natural antisense mRNA transcript of MEAC1 and MEAC2 from neonatal mouse brain, complete cds.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF087578; AAC61495.1; -;
SQ SEQUENCE 76 AA; 7857 MW; DD4C6EBCED5029F7 CRC64;

Query Match 100.0%; Score 20; DB 11; Length 76;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 41 AAPM 44

RESULT 15
Q8S7E4 PRELIMINARY; PRT; 77 AA.
ID Q8S7E4;
AC Q8S7E4;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Hypothetical 8.1 kDa protein.
GN OSJNBA0057L21.13.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N., Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsirlin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., Vanaken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J., Salzberg S.L., White O., Fraser C.M.;
RA "Oryza sativa chromosome 10 BAC OSJNBA0057L21 genomic sequence.";
RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC087599; AAL79692.1; -;
KW Hypothetical protein.
SQ SEQUENCE 77 AA; 8068 MW; D0CB87D68C03FD03 CRC64;

Query Match 100.0%; Score 20; DB 10; Length 77;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 37 AAPM 40

RESULT 16
Q94188 PRELIMINARY; PRT; 79 AA.
ID Q94188;
AC Q94188;
DT 01-FEB-1997 (TREMBlrel. 02, Created)

DT 01-FEB-1997 (TREMBlrel. 02, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE F22H10.3 protein.
GN F22H10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R., Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Wellstock L., Wilkinson-Sproat J., Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";
RT Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Langston, Y., Hawkins J.;
RT "The sequence of C. elegans cosmid F22H10.";
RN Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U70845; AAB09098.1; -;
SQ SEQUENCE 79 AA; 8863 MW; B37353A35B920CB0 CRC64;

Query Match 100.0%; Score 20; DB 5; Length 79;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 47 AAPM 50

RESULT 17
Q91ML0 PRELIMINARY; PRT; 83 AA.
ID Q91ML0;
AC Q91ML0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Major core protein (Fragment).
OS Reovirus sp. RVA.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=165101;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RVA;
RX MEDLINE=21318688; PubMed=11425715;
RA Spinner M.L., Di Giovanni G.D.;
RT "Detection and identification of mammalian reoviruses in surface water by combined cell culture and reverse transcription-pcr.";
RL Appl. Environ. Microbiol. 67:3016-3020(2001).
DR EMBL: AF325764; AAK71927.1; -;
FT NON_TER 1
FT NON_TER 83
FT NON_TER 83
SQ SEQUENCE 83 AA; 9794 MW; 74CD2CCF292B3D50 CRC64;

Query Match 100.0%; Score 20; DB 12; Length 83;

Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 23 AAPM 26

RESULT 18

Q91MK9 PRELIMINARY; PRT; 83 AA.
AC Q91MK9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Major core protein (Fragment).
OS Reovirus sp. RVC.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=165102;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RVC;
RX MEDLINE=21318688; PubMed=11425715;
RA Spinner M.L., Di Giovanni G.D.;
RT "Detection and identification of mammalian reoviruses in surface water
by combined cell culture and reverse transcription-pcr.";
RL Appl. Environ. Microbiol. 67:3016-3020(2001).
DR EMBL; AF325765; AAK71928.1; -.
FT NON_TER 1 1
FT 83 83
SQ SEQUENCE 83 AA; 9794 MW; 74CD2CCF292B3D50 CRC64;

Query Match 100.0%; Score 20; DB 12; Length 83;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 23 AAPM 26

RESULT 19

Q91MK8 PRELIMINARY; PRT; 83 AA.
AC Q91MK8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Major core protein (Fragment).
OS Reovirus sp. RVE.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=165103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RVE;
RX MEDLINE=21318688; PubMed=11425715;
RA Spinner M.L., Di Giovanni G.D.;
RT "Detection and identification of mammalian reoviruses in surface water
by combined cell culture and reverse transcription-pcr.";
RL Appl. Environ. Microbiol. 67:3016-3020(2001).
DR EMBL; AF325766; AAK71929.1; -.
FT NON_TER 1 1
FT 83 83
SQ SEQUENCE 83 AA; 9794 MW; 74CD2CCF292B3D50 CRC64;

Query Match 100.0%; Score 20; DB 12; Length 83;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 23 AAPM 26

RESULT 20

Q91MK7 PRELIMINARY; PRT; 83 AA.
AC Q91MK7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Major core protein (Fragment).
OS Reovirus sp. RVG.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=165104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RVG;
RX MEDLINE=21318688; PubMed=11425715;
RA Spinner M.L., Di Giovanni G.D.;
RT "Detection and identification of mammalian reoviruses in surface water
by combined cell culture and reverse transcription-pcr.";
RL Appl. Environ. Microbiol. 67:3016-3020(2001).
DR EMBL; AF325767; AAK71930.1; -.
FT NON_TER 1 1
FT 83 83
SQ SEQUENCE 83 AA; 9793 MW; 2ED4046DABFDE62C CRC64;

Query Match 100.0%; Score 20; DB 12; Length 83;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 23 AAPM 26

RESULT 21

Q91MK6 PRELIMINARY; PRT; 83 AA.
AC Q91MK6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Major core protein (Fragment).
OS Reovirus sp. RVH.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=165105;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RVH;
RX MEDLINE=21318688; PubMed=11425715;
RA Spinner M.L., Di Giovanni G.D.;
RT "Detection and identification of mammalian reoviruses in surface water
by combined cell culture and reverse transcription-pcr.";
RL Appl. Environ. Microbiol. 67:3016-3020(2001).
DR EMBL; AF325768; AAK71931.1; -.
FT NON_TER 1 1
FT 83 83
SQ SEQUENCE 83 AA; 9794 MW; 74CD2CCF292B3D50 CRC64;

Query Match 100.0%; Score 20; DB 12; Length 83;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
1111
Db 23 AAPM 26

RESULT 22

Q54062 PRELIMINARY; PRT; 83 AA.
ID Q54062;
AC Q54062;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)

```
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Exst protein (Putative transcriptional regulator protein).
GN EXST OR RB1058 OR SMB20935.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RX MEDLINE=96133689; PubMed=8544814;
RA Becker A., Kuester H., Niehaus K., Puehler A.;
RT "Extension of the Rhizobium meliloti succinoglycan biosynthesis gene
RT cluster: identification of the exsA gene encoding an ABC transporter
RT protein, and the exsB gene which probably codes for a . . . .";
RL Mol. Gen. Genet. 249:487-497(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RA York G.M., Walker G.C.;
RT "The Rhizobium meliloti exoK gene and prsD/prsE/exsH genes are
RT components of independent degradative pathways which contribute to
RT production of low-molecular-weight succinoglycan.";
RL Mol. Microbiol. 25:117-134(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RA Becker A.;
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=1021; PLASMID=PSYMB (MEGAPLASMID 2);
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoeelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AJ225561; CAA12534.1; -.
DR EMBL; AL603645; CAA49458.1; -.
DR InterPro; IPR001387; HTH_3.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SM00530; HTH_XRE; 1.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 83 AA; 8681 MW; 80E4E2D00085EFBF CRC64;

Query Match 100.0%; Score 20; DB 16; Length 83;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 13 AAPM 16

RESULT 23
086845 PRELIMINARY; PRT; 84 AA.
AC 086845;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative 38.9 kDa protein (Fragment).
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Palframan W.;
```

```
RL Thesis (1998), University of East Anglia, Department of Biological.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=20402106; PubMed=10931278;
RA Molle V., Buttner M.J.;
RT "Different alleles of the response regulator gene bldM arrest
RT Streptomyces coelicolor development at distinct stages.";
RL Mol. Microbiol. 36:1265-1278(2000).
DR EMBL; AJ010601; CAA09266.1; -.
DR InterPro; IPR001093; IMPDH/GMPrtase.
DR Pfam; PF01574; IMPDH_N; 1.
FT NON TER 84
SQ SEQUENCE 84 AA; 9284 MW; 43FA1467EEB4ED8F CRC64;

Query Match 100.0%; Score 20; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 52 AAPM 55

RESULT 24
0805M7 PRELIMINARY; PRT; 88 AA.
AC 0805M7;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE AGR_C_597P.
GN AGR_C_597.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE007972; AAK86159.1; -.
SQ SEQUENCE 88 AA; 10016 MW; 1919F5932DE43263 CRC64;

Query Match 100.0%; Score 20; DB 16; Length 88;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 64 AAPM 67

RESULT 25
090JAI PRELIMINARY; PRT; 89 AA.
AC 090JAI;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VPI (Fragment).
GN 1D.
OS Foot-and-mouth disease virus O.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
```

```
OX NCBI_TaxID=12118;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUN/1/99;
RA Samuel A.R., Knowles N.J.;
RT "Foot-and-mouth disease type O viruses exhibit genetically and
RT geographically distinct evolutionary lineages (topotypes).";
RL J. Gen. Virol. 82:0-0(2001).
DR EMBL; AJ296323; CAC22166.1; -.
FT NON_TER 1
FT NON_TER 89
SQ SEQUENCE 89 AA; 9916 MW; 9A9C9AE97A0AF468 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 31 AAPM 34
```

```
RESULT 26
O96UH7 PRELIMINARY; PRT; 92 AA.
AC O96UH7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Fructose-1,6-bisphosphate aldolase (Fragment).
OS Paracoccidioides brasiliensis.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Onygenales; mitosporic Onygenales; Paracoccidioides.
OX NCBI_TaxID=121759;
RN [1]
RP SEQUENCE FROM N.A.
RA Carneiro L.C., Jesuino R.S.A., Felipe M.S.S., Pereira M.,
RA Soares C.M.A.;
RT "The fructose-1,6-bisphosphate aldolase genomic sequence of
RT Paracoccidioides brasiliensis.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY057387; AAL25625.1; -.
DR InterPro; IPR000771; F_bp_aldolase.
DR Pfam; PF01116; F_bp_aldolase; 1.
DR ProDom; PD002376; F_bp_aldolase; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 92 AA; 10652 MW; 9B58BDD56F511121 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 3; Length 92;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 71 AAPM 74
```

```
RESULT 27
O9RJE9 PRELIMINARY; PRT; 94 AA.
AC O9RJE9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein SC00743.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
```

```
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL13171; CAB61536.2; -.
KW Hypothetical protein.
SQ SEQUENCE 94 AA; 9520 MW; B252700FDAE10CB1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 16; Length 94;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 21 AAPM 24
```

```
RESULT 28
O47762 PRELIMINARY; PRT; 98 AA.
AC O47762;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NADH ubiquinone oxidoreductase subunit 4L.
GN ND4L.
OS Myxocyprinus asiaticus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Catostomidae; Myxocyprinus.
OX NCBI_TaxID=70543;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21097357; PubMed=11161753;
RA Xiao W., Zhang Y., Liu H.;
RT "Molecular Systematics of Xenocyprinae (Teleostei: Cyprinidae):
RT Taxonomy, Biogeography, and Coevolution of a Special Group Restricted
RT in East Asia.";
RL Mol. Phylogenet. Evol. 18:163-173(2001).
DR EMBL; AF036180; AAD10068.1; -.
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; Oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion; Ubiquinone.
SQ SEQUENCE 98 AA; 10485 MW; F5DBDDA294839E3D CRC64;
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 8; Length 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
Db 58 AAPM 61

RESULT 29
O47775 PRELIMINARY; PRT; 98 AA.
AC O47775;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NADH ubiquinone oxidoreductase subunit 4L.
```


GN ND4L.
OS Xenocypris fangl.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Xenocypris.
OX NCBI_TaxID=70547;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=21097357; PubMed=11161753;
RA Xiao W., Zhang Y., Liu H.;
RT "Molecular Systematics of Xenocyprinae (Teleostei: Cyprinidae):
RT Taxonomy, Biogeography, and Coevolution of a Special Group Restricted
RT in East Asia.";
RL Mol. Phylogenet. Evol. 18:163-173(2001).
DR EMBL: AF036189; AAD10086.1; -.
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam: PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion; Ubiquinone.
SQ SEQUENCE 98 AA; 10543 MW; 0BC09218C5C7565A CRC64;

Query Match 100.0%; Score 20; DB 8; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
| | | |
Db 58 AAPM 61

RESULT 30
047781 PRELIMINARY; PRT; 98 AA.
ID 047781;
AC 047781;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE NADH ubiquinone oxidoreductase subunit 4L.
GN ND4L.
OS Xenocypris yunnanensis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Xenocypris.
OX NCBI_TaxID=70549;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RX MEDLINE=21097357; PubMed=11161753;
RA Xiao W., Zhang Y., Liu H.;
RT "Molecular Systematics of Xenocyprinae (Teleostei: Cyprinidae):
RT Taxonomy, Biogeography, and Coevolution of a Special Group Restricted
RT in East Asia.";
RL Mol. Phylogenet. Evol. 18:163-173(2001).
DR EMBL: AF036193; AAD10094.1; -.
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam: PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion; Ubiquinone.
SQ SEQUENCE 98 AA; 10679 MW; F2D2265885938A2C CRC64;

Query Match 100.0%; Score 20; DB 8; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
| | | |
Db 58 AAPM 61

RESULT 31
09G6S9 PRELIMINARY; PRT; 98 AA.
ID 09G6S9
AC 09G6S9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE NADH dehydrogenase subunit 4L.
GN ND4L.
OS Sardnops melanostictus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;
OC Sardnops.
OX NCBI_TaxID=41697;
RN [1]
RP SEQUENCE FROM N.A.
RA Inoue J.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Inoue J.G., Miya M., Tsukamoto K., Nishida M.;
RT "Complete mitochondrial DNA sequence of the Japanese sardine Sardnops
RT melanostictus.";
RL Fisheries Sci. 66:924-932(2000).
DR EMBL: AB032554; BAB18610.1; -.
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam: PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10482 MW; 226D411ECBC4E709 CRC64;

Query Match 100.0%; Score 20; DB 8; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
| | | |
Db 58 AAPM 61

RESULT 32
09T919 PRELIMINARY; PRT; 98 AA.
ID 09T919
AC 09T919;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE NADH ubiquinone oxidoreductase subunit 4L.
GN ND4L.
OS Aristichthys nobilis (bighead carp).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Aristichthys.
OX NCBI_TaxID=75338;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Xiao W., Zhang Y.;
RT "Phylogenetic analysis of the family Cyprinidae (Teleostei,
RT Cypriniformes) based on complete cytochrome b and ND4L-ND4 sequences:
RT implications for its origination and East Asia biogeography.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF068319; AAF21263.1; -.
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam: PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion; Ubiquinone.
SQ SEQUENCE 98 AA; 10529 MW; 4434C35DC5D0E081 CRC64;

Query Match 100.0%; Score 20; DB 8; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
||||
Db 58 AAPM 61

RESULT 33
Q9T911
ID Q9T911 PRELIMINARY; PRT; 98 AA.
AC Q9T911;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NADH ubiquinone oxidoreductase subunit 4L.
GN ND4L.
OS Megalobrama pellegrini.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Megalobrama.
OX NCBI_TaxID=77646;
RN [1]
RP SEQUENCE FROM N.A.
RA Xiao W., Zhang Y.;
RT "Phylogenetic analysis of the family Cyprinidae (Teleostei,
RT Cypriniformes) based on complete cytochrome b and ND4L-ND4 sequences:
RT Implications for its origin and East Asia biogeography.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF068325; AAF21275.1; -
DR InterPro: IPR001133; Oxidored_4L.
DR InterPro: IPR003214; Oxidred4L.
DR Pfam: PF00420; oxidored_q2; 1.
DR ProDom: PD000359; Oxidred4L; 1.
KW Mitochondrion; Ubiquinone.
SQ SEQUENCE 98 AA; 10529 MW; 442092089495E481 CRC64;

Query Match 100.0%; Score 20; DB 8; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
||||
Db 58 AAPM 61

RESULT 34
Q9MIA1
ID Q9MIA1 PRELIMINARY; PRT; 98 AA.
AC Q9MIA1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NADH dehydrogenase subunit 4L.
GN NADH4L.
OS Typhlonectes natans (Rubber eel).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Gymnophiona; Caeciliidae; Typhlonectes.
OX NCBI_TaxID=8456;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20296752; PubMed=10835397;
RA Zardoya R., Meyer A.;
RT "Mitochondrial Evidence on the Phylogenetic Position of Caecilians
RT (Amphibia: Gymnophiona).";
RL Genetics 155:765-775(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Zardoya R.;

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF154051; AAF78155.1; -
DR InterPro: IPR001133; Oxidored_4L.
DR InterPro: IPR003214; Oxidred4L.
DR Pfam: PF00420; oxidored_q2; 1.
DR ProDom: PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10537 MW; 3EB0BDFD3507A6A9 CRC64;

Query Match 100.0%; Score 20; DB 8; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
||||
Db 58 AAPM 61

RESULT 35
Q9G3Z2
ID Q9G3Z2 PRELIMINARY; PRT; 98 AA.
AC Q9G3Z2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NADH dehydrogenase subunit 4L.
GN NADH-4L.
OS Cottus bairdii (mottled sculpin).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Scorpaeniformes;
OC Cottidae; Cottidae; Cottus.
OX NCBI_TaxID=147208;
RN [1]
RP SEQUENCE FROM N.A.
RA Nelson S.E., Schweny S.P., Baker R.L., Eckdahl T.T.;
RT "Identification of Cottus Species Using Mitochondrial RFLP Analysis.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF312388; AAG47826.1; -
DR InterPro: IPR001133; Oxidored_4L.
DR InterPro: IPR003214; Oxidred4L.
DR Pfam: PF00420; oxidored_q2; 1.
DR ProDom: PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10391 MW; 0766B08C1563C3C4 CRC64;

Query Match 100.0%; Score 20; DB 8; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
||||
Db 58 AAPM 61

RESULT 36
Q9G3Y9
ID Q9G3Y9 PRELIMINARY; PRT; 98 AA.
AC Q9G3Y9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NADH dehydrogenase subunit 4L.
GN NADH-4L.
OS Cottus caroliniae (banded sculpin).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Scorpaeniformes;
OC Cottidae; Cottidae; Cottus.
OX NCBI_TaxID=147209;
RN [1]

```
RP SEQUENCE FROM N.A.
RA Nelson S.E., Schwery S.P., Baker R.L., Eckdahl T.T.;
RT "Identification of Cottus Species Using Mitochondrial RFLP Analysis.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF312389; AAG47829.1; .
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidored4L.
DR Pfam; PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10416 MW; 199AAC82E57D3543 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 8; Length 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 58 AAPM 61

RESULT 37
Q9G3Y6 PRELIMINARY; PRT; 98 AA.
ID Q9G3Y6
AC Q9G3Y6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NADH dehydrogenase subunit 4L.
GN NADH-4L.
OS Cottus hypselurus (Ozark sculpin).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Scorpaeniformes;
OC Cottidae; Cottidae; Cottus.
OX NCBI_TaxID=147211;
RN [1]
RP SEQUENCE FROM N.A.
RA Nelson S.E., Schwery S.P., Baker R.L., Eckdahl T.T.;
RT "Identification of Cottus Species Using Mitochondrial RFLP Analysis.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF312390; AAG47832.1; .
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidored4L.
DR Pfam; PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10421 MW; A77A57D79F4B7ADE CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 8; Length 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 58 AAPM 61

RESULT 38
Q9G3Y3 PRELIMINARY; PRT; 98 AA.
ID Q9G3Y3
AC Q9G3Y3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NADH dehydrogenase subunit 4L.
GN NADH-4L.
OS Cottus cognatus (slimy sculpin).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Scorpaeniformes;
OC Cottidae; Cottidae; Cottus.
OX NCBI_TaxID=147210;
RN [1]
RP SEQUENCE FROM N.A.
RA Baker R.L., Chandler M.L., Eckdahl T.T.;
RT "Identification of Cottus Species in Montana Using Mitochondrial RFLP Analysis.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF312391; AAG47835.1; .
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidored4L.
DR Pfam; PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10410 MW; 077A428C1563C3C4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 8; Length 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 58 AAPM 61

RESULT 39
Q94YP8 PRELIMINARY; PRT; 98 AA.
ID Q94YP8
AC Q94YP8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NADH dehydrogenase subunit 4L.
GN ND4L.
OS Osteoglossum bicirrhosum (silver arawana).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxID=109271;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21370231; PubMed=11476635;
RX Inoue J.G., Miya M., Tsukamoto K., Nishida M.;
RT "A Mitogenomic Perspective on the Basal Teleostean Phylogeny:
RT Resolving Higher-Level Relationships with Longer DNA Sequences.";
RT Mol. Phylogenet. Evol. 20:275-285(2001).
DR EMBL; AB043025; BAB64385.1; .
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidored4L.
DR Pfam; PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10539 MW; 4D6EABC7754E1753 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 8; Length 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPM 4
Db 58 AAPM 61

RESULT 40
Q94TB1 PRELIMINARY; PRT; 98 AA.
ID Q94TB1
AC Q94TB1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NADH dehydrogenase subunit 4L.
GN ND4L.
```

OS Neoscopelus microchir.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Scopelomorpha; Myctophiformes; Neoscopelidae; Neoscopelus.
OX NCBI_TaxID=143321;
RN [1]
RP SEQUENCE FROM N.A.
RA Miya M.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21519020; PubMed=11606696;
RA Miya M., Kawaguchi A., Nishida M.;
RT "Mitogenomic exploration of higher teleostean phylogenies: A case
RT study for moderate-scale evolutionary genomics with 38 newly
RT determined complete mitochondrial DNA sequences."
RL Mol. Biol. Evol. 18:1993-2009(2001).
DR EMBL; AP002921; BAB70061.1; -.
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidored4L.
DR Pfam; PF00420; Oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10368 MW; EB8ED12057A684CB CRC64;

Query Match 100.0%; Score 20; DB 8; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPM 4
 ||||
Db 58 AAPM 61

Search completed: December 6, 2002, 13:30:08
Job time : 19.5556 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:19:54 ; Search time 26 Seconds
(Without alignments)
20.500 Million cell updates/sec

Title: US-10-033-526-1
Perfect score: 21
Sequence: 1 AAPF 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 21 | 100.0 | 4 | 5 | AAP40459 | N-terminal polyhyd |
| 2 | 21 | 100.0 | 4 | 11 | AAR04012 | Peptidase substrat |
| 3 | 21 | 100.0 | 4 | 12 | AAR10887 | Peptide component |
| 4 | 21 | 100.0 | 4 | 12 | AAR11223 | Ketone analogue pr |
| 5 | 21 | 100.0 | 4 | 13 | AAR29964 | Cathepsin G inhibi |
| 6 | 21 | 100.0 | 4 | 14 | AAR34258 | Chromogenic peptid |
| 7 | 21 | 100.0 | 4 | 14 | AAR38410 | Cathepsin G inhibi |
| 8 | 21 | 100.0 | 4 | 14 | AAR44109 | Aminonaphthalene p |
| 9 | 21 | 100.0 | 4 | 15 | AAR52024 | Bacillus alkali pr |
| 10 | 21 | 100.0 | 4 | 15 | AAR46223 | Serine protease in |

| | | | | | | |
|----|----|-------|---|----|----------|--------------------|
| 11 | 21 | 100.0 | 4 | 15 | AAR53781 | Sequence of intern |
| 12 | 21 | 100.0 | 4 | 16 | AAR72920 | Substrate for pept |
| 13 | 21 | 100.0 | 4 | 16 | AAR72876 | Substrate for pept |
| 14 | 21 | 100.0 | 4 | 16 | AAR77315 | Porphyromonas ging |
| 15 | 21 | 100.0 | 4 | 16 | AAR77196 | Cell proliferation |
| 16 | 21 | 100.0 | 4 | 17 | AAR85709 | Degradable peptide |
| 17 | 21 | 100.0 | 4 | 18 | AAW52606 | Serine protease-in |
| 18 | 21 | 100.0 | 4 | 18 | AAW52606 | Substrate peptide |
| 19 | 21 | 100.0 | 4 | 18 | AAW24567 | Substrate for seri |
| 20 | 21 | 100.0 | 4 | 18 | AAW08165 | Proteinase site of |
| 21 | 21 | 100.0 | 4 | 18 | AAW12810 | Synthetic substrat |
| 22 | 21 | 100.0 | 4 | 19 | AAW79700 | B. subtilis subtil |
| 23 | 21 | 100.0 | 4 | 19 | AAW76694 | B. subtilis subtil |
| 24 | 21 | 100.0 | 4 | 19 | AAW61357 | Synthetic oligopep |
| 25 | 21 | 100.0 | 4 | 19 | AAW51451 | Indicator for dete |
| 26 | 21 | 100.0 | 4 | 19 | AAW51610 | Peptide conjugated |
| 27 | 21 | 100.0 | 4 | 20 | AAW48424 | Alkaline protease- |
| 28 | 21 | 100.0 | 4 | 20 | ABB07072 | Alkaline protease |
| 29 | 21 | 100.0 | 4 | 20 | AAW84189 | Peptide comprising |
| 30 | 21 | 100.0 | 4 | 21 | AAW20774 | Chymotrypsin enzym |
| 31 | 21 | 100.0 | 4 | 21 | AAW03092 | Substrate peptide |
| 32 | 21 | 100.0 | 4 | 21 | AAW01922 | Synthetic PIase s |
| 33 | 21 | 100.0 | 4 | 21 | AAW91025 | Model substrate pe |
| 34 | 21 | 100.0 | 4 | 21 | AAW97813 | Protease inhibitor |
| 35 | 21 | 100.0 | 4 | 21 | AAW80465 | Chymotrypsin subst |
| 36 | 21 | 100.0 | 4 | 21 | AAW76808 | Subtilisin substra |
| 37 | 21 | 100.0 | 4 | 21 | AAW78789 | Cathepsin G peptid |
| 38 | 21 | 100.0 | 4 | 21 | AAW49452 | Peptidylprolyl cis |
| 39 | 21 | 100.0 | 4 | 21 | AAW59632 | zTryp3 substrate u |
| 40 | 21 | 100.0 | 4 | 22 | AAU07699 | Peptide substrate |
| 41 | 21 | 100.0 | 4 | 22 | AAW82792 | Model peptide subs |
| 42 | 21 | 100.0 | 4 | 22 | AAW64601 | Artificial substra |
| 43 | 21 | 100.0 | 4 | 22 | AAW64496 | Human peptide #2, |
| 44 | 21 | 100.0 | 4 | 22 | AAW04342 | Peptide #2. Unide |
| 45 | 21 | 100.0 | 4 | 22 | AAW98623 | |

ALIGNMENTS

RESULT 1
AAP40459
ID AAP40459 standard; Protein; 4 AA.
XX
AC AAP40459;
XX
DT 27-NOV-1991 (first entry)
XX
DE N-terminal polyhydroxyalkanoyl peptide.
XX
KW Polyhydroxyalkanoyl peptide; protease; peptidase; prodrug.
XX
PN EP126685-A.
XX
PD 28-NOV-1985.
XX
PF 15-MAY-1984; 84EP-0400984.
XX
PR 16-MAY-1983; 83FR-0308051.
XX
PA (CNRS) CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE.
XX
PI Monsigny M, Mayer R;
XX
DR WPI; 1984-296065/48.
XX
PT N-terminal polyhydroxyalkanoyl peptide and aminoacid derivs. - opt.
PT having C-terminal active gps., e.g. chloroquine, are water-soluble
PT antitumour or antiparasitic cpds. and protease targets.
XX
PS Claim 11; page 15; 20pp; french.
XX
CC The peptide is polyhydroxyalkanoyl-L-Ala-L-Ala-L-Pro-L-Phe-X. X is

CC R2-NH- derived from an aromatic amine, eg P-nitro-aniline, alpha-
CC or beta-naphthylamine, or it is R2NH derived from an amino drug, eg
CC daunorubicin or chloroquine, or it is a gp. which may confer
CC activity, eh H, OH, CH2Cl. When X is the residue of a drug, the
CC cpd. is a prodrug in which the peptide is the substrate for a
CC specific protease secreted by the target cell for the active drug,
CC eg tumour cells or microorganism pathogens, eg Plasmodium
CC falciparum. Due to the polyhydroxy gp., the cpds. are very soluble
CC in water or buffer soln. and are able to bu used at high concn. The
CC cpds. also allow the detection of proteases and peptidases and
CC allow easy determination of the best substrate for a particular
CC protease.
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 2
AAR04012
ID AAR04012 standard; peptide; 4 AA.
XX
AC AAR04012;
XX

DT 19-NOV-1991 (first entry)
XX

DE Peptidase substrate analogues having peptidase inhibition activity.
XX

KW Elastase; plasminogen activator; cathepsin; calpain; enkephalinase;
KW Angiotensin converting enzyme.
XX

OS Synthetic.
XX

FT Key Location/Qualifiers
FT Modified-site 4 /label= Terminal modified from -COOH to -COCOR
FT /note= "R = alkyl group or -H"
XX

PN EP363284-A.
XX

PD 11-APR-1990.
XX

PF 06-OCT-1989; 89EP-0402762.
XX

PR 07-OCT-1988; 88US-0254762.
PR 06-OCT-1989; 89EP-0402763.
XX

PA (RICH) MERRELL DOW PHARM INC.
XX

PI Bey P, Angelastro M, Mehdi S;
XX

DR WPI; 1990-109579/15.
XX

PT New peptidase substrate analogue cpds. - useful as protease
PT inhibitors in treatment of disease states.
XX

PS Claim 5; page 26; 33pp; English.
XX

CC The analogues may be useful in treatment of a variety of disease
CC states. The scissile amide group is replaced with H or a substituted
CC Carbon moiety effectively inhibiting the activity of peptidases such
CC as elastase, plasmin thrombin, urokinase etc.
XX

SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 11; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 3
AAR10887
ID AAR10887 standard; Protein; 4 AA.
XX
AC AAR10887;
XX

DT 10-APR-1991 (first entry)
XX

DE Peptide component of pentafluoroethylcarbonyl analogue.
XX

KW Protease inhibitor; rheumatoid arthritis; thrombosis; psoriasis;
KW male contraceptive.
XX

OS Synthetic.
XX

PN EP410411-A.
XX

PD 30-JAN-1991.
XX

PF 25-JUL-1990; 90EP-0114250.
XX

PR 26-JUL-1989; 89US-0385624.
XX

PA (RICH) MERRELL DOW PHARM INC.
XX

PI Bey P, Peet NP, Angelastro MR, Mehdi S;
XX

DR WPI; 1991-030811/05.
XX

PT Novel serine-, carboxylic acid- and metallo-proteinase-inhibitors
PT - inhibit range of proteinase(s) in treating rheumatoid
PT arthritis, thrombosis and psoriasis, also is a male contraceptive
XX
PS Claim 5; page 25; 40pp; English.
XX

CC This peptide is the R1 gp. of the cpd. of formula: R1NHCH(R2)-
CC COCF2CF3. This cpd. is a peptidase substrate analogue and is a
CC specific enzyme inhibitor for a range of proteases, e.g. serine-,
CC carboxylic acid-, and metallo-proteases. It is useful in the
CC treatment of rheumatoid arthritis, thrombosis and psoriasis and is
CC also used as a male contraceptive. See also AAR10876-83, AAR10886 and
CC AAR10888.
XX

SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 4
AAR11223
ID AAR11223 standard; Protein; 4 AA.
XX

AC AAR11223;
XX

DT 24-MAY-1991 (first entry)
XX

DE ketone analogue protease inhibitor #2.
XX

KW protease inhibitor; antiinflammatory agent; hypotensive; analgesic;
KW antiproliferative agent; antidemyelinating agent; antithrombotic.
XX

OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site /label= Methoxysuccinyl-Ala
FT Modified-site 4
FT /label= Phe-carboxyoxalyl
XX
PN EP417721-A.
XX
PD 20-MAR-1991.
XX
PF 11-SEP-1990; 90EP-0117461.
XX
PR 11-SEP-1989; 89US-0405491.
XX
PA (RICH) MERRELL DOW PHARM INC.
XX
PI Flynn GA, Bey P;
XX
DR WPI; 1991-081980/12.
XX
PT New ketone analogue peptidase and isomerase inhibitors - for
PT inhibition of leukocyte elastase, cathepsin G, thrombin,
PT chymotrypsin, plasmin etc.
XX
PS Claim 5; Page 26; 50pp; English.
XX
CC This peptide is a specific example of a highly generic protease
CC inhibitor useful for medical purposes. The peptide analogues include
CC inhibitors of urokinase, renin, cathepsin D, etc. which can be used
CC as anti-proliferative agents and abortifacients, hypotensives,
CC antiinflammatory and antidemyelinating agents, respectively.
CC See also AAR11222 and AAR11224-R11238.
XX
SQ Sequence 4 AA;
OY 1 AAPF 4
Db 1 AAPF 4
Query Match 100.0%; Score 21; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 5
AAR29964
ID AAR29964 standard; peptide; 4 AA.
XX
AC AAR29964;
XX
DT 19-APR-1993 (first entry)
XX
DE Cathepsin G inhibiting fragment.
XX
KW Cathepsin G; elastase; connective tissue; degradation; protease;
KW gout; rheumatoid arthritis; emphysema; ARDS;
KW adult respiratory distress syndrome; para-phenylene.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site /note= "N-terminal amino acids of the peptides of
FT AAR29963-64 are pref. linked by
FT -C(O)-phenylene-C(O)-", esp. wherein the
FT phenylene is a para-phenylene gp."
FT Modified-site 4
FT /note= "C-terminal Phe is in keto form, i.e. OH
FT replaced by CGIM, pref. CF3 or CF2CF3"
XX
PN WO9220357-A.

XX
PD 26-NOV-1992.
XX
PF 21-APR-1992; 92WO-US03288.
XX
PR 23-MAY-1991; 91US-0704449.
XX
PA (RICH) MERRELL DOW PHARM INC.
XX
PI Angelastro MR, Bey P, Doherty NS, Janusz MJ, Mehdi S;
PI Peet NP;
XX
DR WPI; 1992-415461/50.
XX
PT New peptide derivs. used as cathepsin G and elastase inhibitors -
PT for treating gout, rheumatoid arthritis, inflammatory disorders,
PT emphysema and adult respiratory distress syndrome
XX
PS Claim 13-17; Page 52; 53pp; English.
XX
CC This sequence is an example of a highly generic formula.
CC Inhibitors of cathepsin G and elastase for preventing connective
CC tissue degradation are chemically linked inhibitors of the proteases
CC elastase (pref. the peptide of AAR29963, or Lys(2CBz)-Pro-Val or
CC Val-Pro-Val) and cathepsin G (pref. the peptide of AAR29964, or
CC Val-Pro-Phe or Phe).
CC The N-terminal amino acids of the peptides of AAR29963-64 are pref.
CC linked by -C(O)-phenylene-C(O)-, esp. wherein the phenylene is a
CC para-phenylene gp.
CC The cpds. have an anti-inflammatory effect useful in the treatment
CC of gout, rheumatoid arthritis and other inflammatory diseases and
CC to prevent elastin mediated tissue damage. They can also be used
CC in the treatment of emphysema and adult respiratory distress syndrome.
XX
SQ Sequence 4 AA;
OY 1 AAPF 4
Db 1 AAPF 4
Query Match 100.0%; Score 21; DB 13; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 6
AAR34258
ID AAR34258 standard; peptide; 4 AA.
XX
AC AAR34258;
XX
DT 19-AUG-1993 (first entry)
XX
DE Chromogenic peptide substrate for peptidase assay.
XX
KW Alkaline protease; detergent; stability.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site /note= "N-succinyl alanine"
FT Modified-site 4
FT /note= "Phe-p-nitroanilide"
XX
PN WO9307276-A.
XX
PD 15-APR-1993.
XX
PF 07-OCT-1992; 92WO-US08341.
XX
PR 08-OCT-1991; 91US-0772087.
XX

PA (CHEM-) CHEMGEN CORP.
PA (VIST-) VISTA CHEMICAL CO.
XX
PI Fodge DW, Hsiao H, Lalonde JJ;
XX
DR WPI; 1993-134465/16.
XX
PT Alkaline protease produced by bacillus stable in alkaline
PT conditions - used in detergents and bleaches to decompose
PT proteinaceous stains
XX
PS Disclosure; Page 8; 58pp; English.
XX
CC The peptidase is a chromogenic substrate for use in a peptidase assay
CC to test the peptidase activity of a novel alkaline protease.
CC Detergent was found to interfere with the assay by reducing the
CC protease activity. See also AAR34259-61 and AAR34463-66.
CC
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 14; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
Db 1 AAPF 4

RESULT 7
AAR38410
ID AAR38410 standard; protein; 4 AA.
XX
AC AAR38410;
XX
DT 29-OCT-1993 (first entry)
XX
DE Cathepsin G inhibitor peptide #1.
XX
KW Inflammatory disease; prevention; rheumatoid arthritis; emphysema;
KW neutrophil-mediated connective tissue degradation; gout; elastase;
KW inhibition; adult respiratory distress syndrome.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Modified-site 1
FT /note= "Linked via a linking chain such as
FT -CO-phenylene-CO- to the N-terminal of a
FT Cathepsin G inhibiting molecule as described
FT in AAR38410-11"
FT Modified-site 4
FT /note= "May opt. be modified by COCOR, CF2CF3, CF3,
FT CHF2, COOR3, CONHR3, CF2CHR3CONHR, H, alkyl,
FT aryl, aralkyl or COR; where R3 is H, alkyl,
FT phenyl or benzyl and R is OH or alkoxy"
XX
PN ZA9203602-A.
XX
XX 24-FEB-1993.
PD
XX
PF 18-MAY-1992; 92ZA-0003602.
XX
XX 23-MAY-1991; 91US-0704499.
PR
XX (RICH) MERRELL DOW PHARM INC.
PA
PI Angelastro MM, Bey P, Doherty NS, Janusz MJ, Mehdi S;
PI Peet NP;
XX WPI; 1993-197380/24.
DR
XX New cathepsin G and elastase inhibitors - prevents connective

PT tissue degradation
XX
XX Claim 13-17; Page 52; 55pp; English.
PS
XX
CC The sequences given in AAR38410-11 are cathepsin G inhibiting peptides
CC which were produced by standard peptide synthesis methods. In the
CC context of the invention one of these peptides may be linked via
CC their N-termini to an elastase inhibiting peptide (see features
CC table). The peptide conjugates may be used to prevent neutrophil-
CC mediated connective tissue degradation associated with inflammatory
CC diseases eg. gout and rheumatoid arthritis. They may also be used
CC for preventing elastin-mediated tissue damage in the treatment of
CC emphysema and adult respiratory distress syndrome.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 14; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
Db 1 AAPF 4

RESULT 8
AAR44109
ID AAR44109 standard; peptide; 4 AA.
XX
AC AAR44109;
XX
DT 16-MAY-1994 (first entry)
XX
DE Aminonaphthalene propyl sulphamide tetrapeptide.
XX
KW chymotrypsin assay; ansa-substrate; detector group.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Modified-site 1
FT /label= OTHER
FT /note= "Suc-Ala (Claim 1) or Z-Ala (Claim 2)"
FT Modified-site 4
FT /label= OTHER
FT /note= "Phe-NH-(1,5-naphthylene)-SO2NH-C3H7"
XX
PN SU1771478-A.
XX
XX 23-OCT-1992.
PD
XX
XX 20-JUL-1990; 90SU-4872552.
PF
XX 20-JUL-1990; 90SU-4872552.
XX
PR 20-JUL-1990; 90SU-4872552.
XX
PA (ALIB-) AS LITH BIOCHEM INST.
PA (ASMO-) AS USSR MOLECULAR GENETICS.
XX
PI Nedospasov AA, Palaima AI, Yanchene RA;
PI
XX WPI; 1993-358242/45.
DR
XX
XX Prepn. of new succinyl-alanyl-alanyl-propyl-phenylalanyl-amino-
PT naphthalen-propyl-sulphamide - for use as ansa-substrate in the
PT quantitative determ. of chymotrypsin
PT
XX Claim 1 and 2; Column 8; 4pp; Russian.
PS
XX
CC 5-(N-succinyl-alanyl-alanyl-propyl-phenylalanyl)aminonaphthalene-1-
CC (N-propyl)sulphamide (Claim 1) and 5-(N-benzoyloxycarbonyl-alanyl-
CC alanyl-propyl-phenylalanyl)aminonaphthalene-1-(N-propyl)sulphamide
CC (Claim 2) are prepared by initially reacting N-benzoyloxycarbonyl-
CC phenylalanine with isobutyl chloroformate in tetrahydrofuran

CC followed by reaction with 5-aminonaphthalene-1-(N-propyl)sulphamide.
CC The new compounds are used as detector groups in quantitative
CC determination of chymotrypsin.

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 14; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
 ||||
Db 1 AAPF 4

RESULT 9
AAR52024

ID AAR52024 standard; peptide; 4 AA.

XX AC AAR52024;

XX DT 28-NOV-1994 (first entry)

XX DE Bacillus alkali protease substrate.

XX KW Bacillus NKS-21 alkali protease; hydrolysis; thermostable enzyme;

XX KW detergent stable.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1

FT /label= OTHER

FT /note= "Suc-Ala"

FT Modified-site 4

FT /label= OTHER

FT /note= "Phe-MCA"

XX PN JP06070765-A.

XX PD 15-MAR-1994.

XX PF 08-OCT-1992; 92JP-0296360.

XX PR 10-JUL-1992; 92JP-0207302.

XX PA (SHOW) SHOWA DENKO KK.

XX DR WPI; 1994-128672/16.

XX PT New alkali protease stable to heat and detergent - useful as

XX PT industrial enzyme, eg in washing compositions

XX PS Claim 1; Page 2; 10pp; Japanese.

CC A novel alkali protease obtained from Bacillus NKS-21 is defined by

CC its physico-chemical properties including substrate specificity for

CC the tetrapeptide AAR52024.

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 15; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
 ||||
Db 1 AAPF 4

RESULT 10
AAR46223

ID AAR46223 standard; peptide; 4 AA.

XX AAR46223;

XX DT 04-AUG-1994 (first entry)

XX DE Serine protease inhibitor tetrapeptide.

XX KW prevention; schistosomiasis; parasite; infection; prevention;

XX KW parasitic penetration; skin; cercariae; anti-penetrant.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "Bg(peptide blocking gp.) attached"

FT Modified-site 4

FT /note= "PI(protease inhibitor), other than

FT chloromethyl ketone, attached"

XX PN US5284829-A.

XX PD 08-FEB-1994.

XX PF 26-NOV-1991; 91US-0798565.

XX PR 26-NOV-1991; 91US-0798565.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Cohen FE, McKerrow JH;

XX DR WPI; 1994-056364/07.

XX PT Synthetic tetra:peptide(s) with an N-terminal blocking gp. and

XX PT C-terminal enzyme inhibitor - can be formulated into soaps and

XX PT sprays and used to prevent schistosomal skin penetration

XX PS Disclosure; Page 7; 35pp; English.

CC The sequence is that of a synthetic tetrapeptide serine protease

CC inhibitor which can be used to prevent schistosome parasite

CC infection. It may be used in a formulation as a soap, lotion,

CC cream, spray, etc. to stop parasitic penetration of the skin.

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 15; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
 ||||
Db 1 AAPF 4

RESULT 11
AAR53781

ID AAR53781 standard; Peptide; 4 AA.

XX AC AAR53781;

XX DT 29-DEC-1994 (first entry)

XX DE Sequence of internal fragment of a chromogenic substrate for

XX DE cathepsin G.

XX KW Enzyme; cathepsin G; protease; chromogenic substrate.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1

FT /label= Suc-Ala

FT Modified-site /note= "Succinyl group"
FT 4
FT /label= Phe-PNA
FT /note= "p-nitroanil"
XX
XX (PN W09412637-A.)
XX
PD 09-JUN-1994.
XX
PF 01-DEC-1993; 93WO-US11696.
XX
PR 02-DEC-1992; 92US-0985692.
PR 19-NOV-1993; 93US-0155331.
XX
PA (NOVO) NOVO-NORDISK AS.
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Foster DC, Norris K, Sprecher CA;
XX
DR WPI; 1994-200265/24.
XX
PT New human Kunitz type protease inhibitor and related DNA - for
PT treating pancreatitis and other disorders involving serine
PT protease, also new amyloid protein precursor homologues including
PT the inhibitor in its sequence
XX
PS Example; Page 63; 70pp; English.
XX
CC Protease inhibitory profiles of the Kunitz inhibitors
CC were determined for a variety of proteases using
CC a variety of chromogenic substrates and compared to the inhibitory
CC activity shown by the Kunitz-type inhibitor domain of the amyloid
CC protein precursor and bovine aprotinin. The substrate AAR53780
CC was tested using the protease leukocyte elastase.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 15; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 1 AAPF 4

RESULT 12
AAR72920
ID AAR72920 standard; Peptide; 4 AA.
XX
AC AAR72920;
XX
DT 29-NOV-1995 (first entry)
XX
DE Substrate for peptidyl prolyl cis trans isomerase alpha.
XX
KW Escherichia coli; protein conformation; folding; acceleration;
KW PPIase-alpha; peptidyl prolyl cis trans isomerase alpha;
KW catalysis; isomerisation; prolyl peptide bond.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Modified-site 1
FT /note= "N-succinyl-Ala"
FT Modified-site 4
FT /note= "Phe-MCA (sic)"
XX
PN EP647714-A.
XX
PD 12-APR-1995.
XX
PF 19-JUL-1990; 90EP-0307914.

XX 19-JUL-1989; 89JP-0184738.
PR 06-OCT-1989; 89JP-0260244.
PR 29-DEC-1989; 89JP-0344705.
XX
XX (TOFU) TONEN CORP.
XX
PI Hayano T, Katou S, Maki N, Suzuki M, Takahashi N;
XX
DR WPI; 1995-140756/19.
XX
PT New E.coli peptidyl prolyl cis trans isomerase beta - used to
PT accelerate the folding of proteins, partic. for activation of
PT inactive recombinant proteins
XX
PS Disclosure; Page 5; 85pp; English.
XX
CC To measure the activity of E. coli PPIase-alpha (peptidyl prolyl cis
CC trans isomerase alpha), the PPIase and CSA (cyclosporine A) are
CC added in the cell and mixed for 1 minute. Then, this peptide is
CC added and incubated, and chymotrypsin added to start the reaction.
CC The effect of CSA on the inhibition of the PPIase activity can be
CC detected by varying the amt. of the CSA added. The inventors are
CC claiming a PPIase-beta.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 16; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 1 AAPF 4

RESULT 13
AAR72876
ID AAR72876 standard; Peptide; 4 AA.
XX
AC AAR72876;
XX
DT 29-NOV-1995 (first entry)
XX
DE Substrate for peptidyl prolyl cis trans isomerase alpha.
XX
KW Escherichia coli; protein conformation; folding; acceleration;
KW PPIase-alpha; peptidyl prolyl cis trans isomerase alpha;
KW catalysis; isomerisation; prolyl peptide bond.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Modified-site 1
FT /note= "N-succinyl-Ala"
FT Modified-site 4
FT /note= "Phe-MCA (sic)"
XX
PN EP647713-A.
XX
PD 12-APR-1995.
XX
PF 19-JUL-1990; 90EP-0307914.
XX
PR 19-JUL-1989; 89JP-0184738.
PR 06-OCT-1989; 89JP-0260244.
PR 29-DEC-1989; 89JP-0344705.
XX
PA (TOFU) TONEN CORP.
XX
PI Hayano T, Katou S, Maki N, Suzuki M, Takahashi N;
XX
DR WPI; 1995-140755/19.


```
XX New E.coli peptidyl prolyl cis trans isomerase alpha - used to
PT accelerate the folding of proteins, partic. for activation of
PT inactive recombinant proteins
XX
PS Disclosure; Page 5; 85pp; English.
XX
CC To measure the activity of E. coli PPIase-alpha (peptidyl prolyl cis
CC trans isomerase alpha), the PPIase and CSA (cyclosporine A) are
CC added in the cell and mixed for 1 minute. Then, this peptide is
CC added and incubated, and chymotrypsin added to start the reaction.
CC The effect of CSA on the inhibition of the PPIase activity can be
CC detected by varying the amt. of the CSA added. The inventors are
CC claiming the PPIase-alpha.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 16; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
   ||||
Db 1 AAPF 4

RESULT 14
AAR77315
ID AAR77315 standard; Protein; 4 AA.
XX
AC AAR77315;
XX
DT 28-FEB-1996 (first entry)
XX
DE Porphyromonas gingivalis protease substrate.
XX
KM Protease; periodontal disease; pathogenic microbe; diagnosis;
KW substrate.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "Suc-Ala"
FT Modified-site 4 /note= "MCA-Phe"
FT
FT
FT
PN JP07135973-A.
XX
PD 30-MAY-1995.
XX
PF 15-NOV-1993; 93JP-0307084.
XX
PR 15-NOV-1993; 93JP-0307084.
XX
PA (SUNR ) SUNTORI,LTD.
XX
DR WPI; 1995-227397/30.
XX
XX
PT An enzyme originated from a periodontal disease pathogenic microbe -
PT and an antibody against the enzyme, for the determination of the
PT progress and activity of the disease
XX
PS Example; Page 7; 15pp; Japanese.
XX
CC AAR77315 is a target substrate for a Porphyromonas gingivalis (a
CC periodontal disease pathogenic microbe) protease. An antibody
CC raised against the enzyme can be used to diagnose the presence
CC and progress of a periodontal disease, caused by a pathogenic
CC microbe.
XX
SQ Sequence 4 AA;
```

```
Query Match 100.0%; Score 21; DB 16; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
   ||||
Db 1 AAPF 4

RESULT 15
AAR77196
ID AAR77196 standard; peptide; 4 AA.
XX
AC AAR77196;
XX
DT 27-FEB-1996 (first entry)
XX
DE Cell proliferation enzyme proteinase activity substrate peptide #1.
XX
KW Proteinase; cell growth-stimulating protein; hydrolysis;
KW macrophage chemotactic action; serine protease inhibitor; wound;
KW gastric ulcer; leg ulcer; bed sore.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /label= Succinyl-Ala
FT Modified-site 4 /note= "Modified by 4-methyl-coumaryl-7-amide"
FT
FT
FT
PN EP661293-A2.
XX
PD 05-JUL-1995.
XX
PF 22-DEC-1994; 94EP-0120406.
XX
PR 29-DEC-1993; 93JP-0351225.
XX
PA (SANW ) SANWA KAGAKU KENKYUSHO CO.
XX
PI Awaya J, Uesaka H, Watanabe K, Yamaguchi T;
XX
DR WPI; 1995-233274/31.
XX
PT Protein derived from Clostridium perfringens FERM BP-4584 -
PT stimulates cell proliferation and has macrophage chemotactic action
XX
PS Example; Page 8; 14pp; English.
XX
CC The sequences given in AAR77196-200 are peptides which were used to
CC demonstrate the proteinase activity of the cell growth-stimulating
CC protein of the invention. The protein was seen to selectively
CC hydrolyse synthetic substrates having an aromatic amino acid at the
CC C-terminal. The response to the peptide given in AAR77196 was
CC particularly high. The cell growth-stimulating protein has a mol.
CC wt. of 420 +/- 40 kD and a single subunit mol. wt. of 130 +/- 20 kD.
CC It has an isoelectric point of 4.8 and has cell growth stimulating
CC action and macrophage chemotactic action, as well as proteinase
CC activity. Its enzymatic activity decreases in the presence of a
CC serine protease inhibitor, increase in the presence of various metal
CC ions, and is stabilised in the presence of calcium ion. The protein is
CC particularly useful for the treatment of wounds, gastric and leg ulcers,
CC eg. bed sores.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 16; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
   ||||
```

Db 1 AAPF 4

RESULT 16
AAR85709
ID AAR85709 standard; peptide; 4 AA.
XX
AC AAR85709;
XX
DT 20-JUN-1996 (first entry)
XX
DE Degradable peptide group, medical polymer gel component.
XX
KW Degradable; polymer gel; wound healing; cover; pseudo-skin; ulcer;
KW bed sore; drug release.
XX
OS Synthetic.
XX
PN WO9531223-A1.
XX
PD 23-NOV-1995.
XX
PE 08-MAY-1995; 95WO-JP00873.
XX
PR 13-MAY-1994; 94JP-0124158.
XX
PA (KURS) KURARAY CO LTD.
XX
PI Kinoshita H, Tanihara M;
XX
DR WPI; 1996-010700/01.
XX
PT Medical polymer gel for wound dressing etc. - comprising
PT water-swellaible gel, spacer, enzyme-hydrolysable unit and active
PT component
XX
PS Claim 3; Page 53; 62pp; Japanese.
XX
CC AAR85708-R85710 are small degradable peptide groups used as components
CC of a medical polymer gel for wound dressing. The gel comprises a
CC water swellaible polymer gel, a spacer molecule, a degradable gp.
CC with a main chain that can be enzymatically broken (as above) and a
CC drug. The gel is heat-resistant, transparent and biocompatible, with
CC high safety. The gel promotes wound healing and may opt. contain,
CC e.g. growth factors, metalloproteinase inhibitors, antibiotics,
CC steroids, etc. The gel is used to cover wounds including cuts, burns
CC and surgical wounds; as a protective cover (pseudo-skin) for bed
CC sores and ulcers; as an adhesive for living tissue; to reinforce bone
CC and as a drug release material.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 17; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
| | | |
Db 1 AAPF 4

RESULT 17
AAW52606
ID AAW52606 standard; peptide; 4 AA.
XX
AC AAW52606;
XX
DT 22-JUN-1998 (first entry)
XX
DE Serine protease-inhibiting peptide with C-terminal phosphonate residue.
XX
KW Selective; serine protease inhibitor; trypsin; elastase; chymotrypsin;
KW antinflammatory; anticoagulant; antitumour.

XX
OS Synthetic.
XX
FH Key
FT Modified-site 1 Location/Qualifiers
FT Modified-site 4 /note= "Meo-Suc-Ala"
FT Modified-site 4
FT /note= "phe(P)-(Oph)2; where phe(P) represents a
FT phenylalanine analogue in which the carboxyl group
FT -COOH has been replaced by a phosphonate group
FT -P(=O)(OH)(OH); and (Oph)2 indicates that the
FT phosphonate has been diphenyl esterified"
XX
PN US5686419-A.
XX
PD 11-NOV-1997.
XX
PF 21-JAN-1994; 94US-0184286.
XX
PR 21-JAN-1994; 94US-0184286.
XX
PA (GEOR-) GEORGIA TECH RES CORP.
XX
PI Boduszek B, Oleksyszyn J, Powers JC;
XX
XX WPI; 1997-558177/51.
XX
PT New peptidyl derivatives of alpha-aminoalkyl phosphonic acid
PT diester - are serine protease inhibitors, useful e.g. for reducing
PT blood coagulation, controlling tumour invasion or treating
PT inflammation
XX
PS Example 14; Column 22; 16pp; English.
XX
CC The patent discloses new peptidyl derivatives of diesters of alpha-
CC aminoalkylphosphonic acids having basic substituents, of formula
CC X-AA4-AA3-AA2-NH-CHR-P(=O)(OZ)(OZ1), in which: R = B-substituted
CC phenyl, B-substituted benzyl or B-substituted 1-6C alkyl; B = amidino,
CC guanidino, isothioureido or amino; Z and Z1 = 1-6C perfluoroalkyl or
CC optionally substituted phenyl; AA2 = an L- or D-amino acid residue;
CC AA3 = a single bond or an L- or D-amino acid residue; AA4 = a single
CC bond or an L- or D-amino acid residue; X = H, NH2CO, NH2CS, NH2SO2,
CC YNHCO, YNHCS, YNHSO2, YCS, YSO2, YOCO, YOCS or YCO; Y = optionally
CC naphthyl, or 1-6C alkyl with 1-2, optionally substituted, attached
CC phenyl groups. The new compounds are inhibitors of trypsin, elastase,
CC chymotrypsin and other serine proteases; and their inhibitory activity
CC is selective depending on the identity of the alpha-aminoalkylphosphonic
CC acid ester residue. They can be used as antinflammatory agents,
CC anticoagulants and antitumour agents. The present sequence is a
CC specific example of the new compounds.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 18; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
| | | |
Db 1 AAPF 4

RESULT 18
AAW30747
ID AAW30747 standard; Protein; 4 AA.
XX
AC AAW30747;
XX
DT 19-JAN-1998 (first entry)
XX
DE Substrate peptide for mutant subtilisin (SAAPF-pna).
XX

| | |
|--|--|
| KW | Subtilisin; BPN'; site specific; protease; substrate assisted catalysis; |
| KW | fusion polypeptide; analogue; variant; mutant; enzyme specificity. |
| XX | |
| OS | Synthetic. |
| XX | |
| FH | Key |
| FT | Modified-site |
| FT | Modified-site |
| FT | Modified-site |
| XX | |
| PN | US5652136-A. |
| XX | |
| PD | 29-JUL-1997. |
| XX | |
| PF | 07-JUN-1995; 95US-0488096. |
| XX | |
| PR | 04-APR-1989; 89US-0334081. |
| PR | 29-MAY-1984; 84US-0614491. |
| PR | 29-MAY-1984; 84US-0614612. |
| PR | 29-MAY-1984; 84US-0614615. |
| PR | 29-MAY-1984; 84US-0614617. |
| PR | 01-APR-1986; 86US-0846627. |
| PR | 30-APR-1986; 86US-0858594. |
| PR | 06-APR-1987; 87US-0035552. |
| PR | 01-DEC-1987; 87US-0127134. |
| PR | 14-JAN-1992; 92US-0823039. |
| PR | 22-SEP-1994; 94US-0287964. |
| PR | 12-JUL-1993; 93US-0090902. |
| XX | |
| PA | (GENEV) GENENCOR INT INC. |
| XX | |
| PI | Carter PJ, Wells JA; |
| XX | |
| DR | WPI; 1997-392947/36. |
| XX | |
| PT | Production of recombinant polypeptide(s) - by expression of fusion |
| PT | polypeptide(s) with a selected target sequence cleavable by a |
| PT | subtilisin-related protease |
| XX | |
| PS | Example 16; Column 41; 44pp; English. |
| XX | |
| CC | This synthetic substrate has been designed to investigate the ability of |
| CC | the subtilisin BPN' analogue S24C:H64A (see AAW30739) to cleave this |
| CC | substrate in the presence of imidazole in order to study substrate |
| CC | assisted catalysis. The specificity of a precursor enzyme is modified by |
| CC | altering or replacing a catalytically functional amino acid at the |
| CC | active site, and also through modification of the target substrate. The |
| CC | target substrate then provides the catalytic activity by interaction |
| CC | with the variant enzyme i.e. the substrate assists in its own catalytic |
| CC | conversion to product. This variant subtilisin, where a catalytic His |
| CC | residue in precursor subtilisin-related protease (wild type) has been |
| CC | replaced by Ala, is capable of cleaving a fusion polypeptide where the |
| CC | target cleavage site is composed of a His residue at position P2 |
| CC | (see AAW30737 for details). The substrate AAPF-pna has no equivalent |
| CC | His residue, however the addition of imidazole has been found to supply |
| CC | the necessary residue. The kcat and Km values for the hydrolysis of |
| CC | variant S24C:H64A subtilisin increased in the presence of imidazole |
| CC | whereas the wild-type enzyme remained unaffected. |
| XX | |
| SQ | Sequence 4 AA; |
| QY | 1 AAPF 4 |
| DB | 1 AAPF 4 |
| Query Match | |
| Best Local Similarity 100.0%; Score 21; DB 18; Length 4; | |
| Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |

RESULT 19
AAW24567

| | |
|--|--|
| ID | AAW24567 standard; peptide; 4 AA. |
| XX | |
| AC | AAW24567; |
| XX | |
| DT | 05-NOV-1997 (first entry) |
| XX | |
| DE | Substrate for serine protease. |
| XX | |
| KW | Serine protease; N-terminus; Streptomyces griseus; guanidine; pre-soak; |
| KW | cleaning composition; laundry detergent; additive composition; enzyme; |
| KW | dishwasher detergent; drain opener; urea; contact lens cleanser; |
| KW | proteolaceous stain. |
| XX | |
| OS | Synthetic. |
| XX | |
| FH | Key |
| FT | Modified-site |
| FT | Modified-site |
| FT | Modified-site |
| XX | |
| PN | US5646028-A. |
| XX | |
| PD | 08-JUL-1997. |
| XX | |
| PF | 18-JUN-1991; 91US-0718303. |
| XX | |
| PR | 18-JUN-1991; 91US-0718303. |
| PR | 06-NOV-1992; 92US-0973343. |
| PR | 18-AUG-1994; 94US-0292924. |
| PR | 17-OCT-1995; 95US-0544143. |
| XX | |
| PA | (CLRX) CLOROX CO. |
| XX | |
| PI | Leigh SD; |
| XX | |
| DR | WPI; 1997-362936/33. |
| XX | |
| PT | Serine protease from Streptomyces griseus ATCC 55178 - with good |
| PT | stability in presence of urea or guanidine, useful in cleaning |
| PT | compositions, including laundry and dishwashing detergents |
| XX | |
| PS | Claim 4; Column 25; 16pp; English. |
| XX | |
| CC | This sequence represents a substrate for the serine protease of the |
| CC | invention. The protease has the N-terminal and C-terminal sequences |
| CC | represented by AAW24565 and AAW24566 respectively. The serine protease |
| CC | was isolated from Streptomyces griseus variety alkaliphilus No. 33 (ATCC |
| CC | 55178). The protease has an apparent molecular weight of 19 kD (by |
| CC | reducing sodium dodecylsulphate polyacrylamide gel electrophoresis), and |
| CC | improved stability against urea and guanidine. The serine protease is |
| CC | specific for the substrate represented by this sequence, but also |
| CC | recognises the substrates shown in AAW26078-W26096. The protease is |
| CC | inhibited by phenylmethylsulphonyl fluoride. The serine protease is |
| CC | useful in liquid or granular cleaning compositions, specifically laundry |
| CC | detergents or additive compositions. It is also useful in automatic |
| CC | dishwasher detergents, pre-soaks, drain openers, contact lens cleansers |
| CC | etc. The protease has better activity against proteolaceous stains than |
| CC | known enzymes and unusually high stability in the presence of chaotropic |
| CC | agents. |
| XX | |
| SQ | Sequence 4 AA; |
| QY | 1 AAPF 4 |
| DB | 1 AAPF 4 |
| Query Match | |
| Best Local Similarity 100.0%; Score 21; DB 18; Length 4; | |
| Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |

RESULT 20

AAW08165
ID AAW08165 standard; Peptide; 4 AA.
XX
AC AAW08165;
XX
DT 26-AUG-1997 (first entry)
XX
DE Proteinase site of a TGF-beta fusion protein.
XX
KW Transforming growth factor-beta fusion protein; wound healing;
KW artificial skin; surgery recovery time.
XX
OS Synthetic.
XX
PN WO9639430-A1.
XX
PD 12-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US08973.
XX
PR 06-JUN-1995; 95US-0470837.
XX
PA (CHEU/) CHEUNG D T.
PA (HALL/) HALL F L.
PA (NIMN/) NIMNI M E.
PA (TUAN/) TUAN T.
PA (WULL/) WU L.
XX
PI Cheung DT, Hall FL, Nimni ME, Tuan T, Wu L;
XX
DR WPI; 1997-043065/04.
XX
PT Prepn. of transforming growth factor-beta fusion protein - useful to
PT reduce surgery recovery time and to prepare artificial skin
XX
PS Claim 8; Page 33; 59pp; English.
XX
CC A novel transforming growth factor-beta (TGF-beta) fusion protein
CC comprises a purification tag and a TGF active fragment. The present
CC sequence represents a specifically claimed proteinase site.
CC Additionally, the fusion protein may comprise proteinase-sensitive
CC linker sites and binding domain so the protein sequence may contain
CC some or all of the following elements: purification tag:proteinase
CC site:ECM binding site:proteinase site:TGF-beta. TGF-beta promotes
CC wound healing, and the fusion protein can be used to reduce surgery
CC recovery time and in the preparation of artificial skin. The inclusion
CC of a purification tag facilitates purification of the fusion protein.
CC The proteinase site is included to permit cleavage and release of the
CC purification tag after purification if desired. The extracellular
CC matrix binding site facilitates delivery of the fusion protein to the
CC desired site of action. Delivery of the TGF-beta to the site to be
CC treated reduces the amount of TGF-beta required to be administered to
CC be effective and reduces the concentration of circulating TGF-beta
CC which may result in undesirable effects.
XX
SQ Sequence 4 AA;
XX
QY 1 AAPF 4
Y 1111
Db 1 AAPF 4
XX
RESULT 21
AAW12810
ID AAW12810 standard; peptide; 4 AA.
XX
AC AAW12810;
XX
DT 21-APR-1997 (first entry)
XX

XX
DE Synthetic substrate #1 for cold-adapted alkali protease.
XX
KW Cold-adapted alkali protease; alteromonas; urea-denatured haemoglobin;
KW yolk; casein; p-nitroaniline; detergent; meat softener.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= Succinylated
FT Modified-site 4
FT /label= p-nitroanilnylated
XX
PN JP08322562-A.
XX
PD 10-DEC-1996.
XX
PF 01-JUN-1995; 95JP-0135056.
XX
PR 01-JUN-1995; 95JP-0135056.
XX
PA (KAOS) KAO CORP.
XX
DR WPI; 1997-081078/08.
XX
PT Cold-adapted alkali protease - produced by Alteromonas, used in
PT detergents
XX
PS Claim 1; Page 2; 10pp; Japanese.
XX
CC AAW12810-W12812 represent synthetic substrates for the cold-adapted
CC alkali protease of the invention. The alkali protease of the invention
CC is produced by Alteromonas, specifically by Alteromonas species KSM-SP
CC 111 which is isolated from a shellfish in the Antarctic ocean. The
CC enzyme of the invention has an active temperature of 0-70 degrees C, with
CC an optimum of 40 degrees, and with 20% of activity retained at 10
CC degrees, and 10% maintained at 0 degrees. The pH range of the enzyme is
CC 4-14, with an optimum of 11, but 70% (or higher) of the activity
CC maintained at pH12. The protease has a molecular weight of 54000, as
CC determined by SDS-polyacrylamide gel electrophoresis. The enzyme is
CC capable of acting on casein, urea-denatured haemoglobin, yolk, and the
CC synthetic substrates represented by these sequences, to yield
CC p-nitroaniline. The enzyme is inhibited by the Hg and Zn metal ions,
CC EDTA, phenylmethanesulphonyl fluoride, chymostatin or
CC p-chloromercuribenzoic acid. Activity of the enzyme is increased 2-7
CC times by the presence of sodium alkanesulphate, sodium
CC alpha-olefinesulphate, sodium polyoxyethylenealkylsulphate, softanol 70H
CC or alpha-sulphofatty acid ester. The protease can be used as a component
CC of detergents, or as a softener for meat when used at a lower
CC temperature.
XX
SQ Sequence 4 AA;
XX
QY 1 AAPF 4
Y 1111
Db 1 AAPF 4
XX
RESULT 22
AAW79700
ID AAW79700 standard; Protein; 4 AA.
XX
AC AAW79700;
XX
DT 25-JAN-1999 (first entry)
XX
DE B. subtilis subtilisin E assay solution peptide.
XX

KW Subtilisin E; mutant; template; in vitro mutagenesis; primer extension.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site /note= "Modified by presence of succinyl group"
FT Modified-site 4
FT Modified-site /note= "Modified by presence of p-nitroanillide"
XX
PN WO9842728-A1.
XX
PD 01-OCT-1998.
XX
PF 25-MAR-1998; 98WO-US05814.
XX
PR 04-AUG-1997; 97US-0905359.
PR 25-MAR-1997; 97US-0041666.
PR 30-APR-1997; 97US-0045211.
PR 12-MAY-1997; 97US-0046256.
XX
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX
PI Affholter JA, Arnold FH, Giver LJ, Shao Z, Zhao H;
XX
DR WPI; 1998-542275/46.
XX
PT New mutagenised polynucleotide encoding ECB deacylase enzyme -
PT produced by the polymerase-catalysed extension of primers
XX
PS Example 6; Page 54; 115pp; English.
XX
CC This peptide is used in a novel method for producing double-stranded
CC mutagenised polynucleotides (PNS) from at least 1 template PN, where the
CC mutagenised PNS have at least 1 nucleotide which is different from the
CC same nucleotide in the template PN. The method is useful for the in vitro
CC mutagenesis and recombination of PNS based on polymerase-catalysed
CC extension of primers. The method can produce full-length genes which are
CC mutants of the original template PNS, and these genes can be further
CC amplified and cloned into vectors. The new PNS can be used in DNA-base
CC computing or for evolving new gene products with improved or new
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 23
AAW76694
ID AAW76694 standard; Peptide; 4 AA.
XX
AC AAW76694;
XX
DT 11-JAN-1999 (first entry)
XX
DE B. subtilis subtilisin E assay solution peptide.
XX
KW Subtilisin E; mutant; template; in vitro mutagenesis; primer extension.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site /note= "Modified by presence of succinyl group"
FT Modified-site 4
FT Modified-site /note= "Modified by presence of p-nitroanillide"

XX
PN WO9842832-A1.
XX
PD 01-OCT-1998.
XX
PF 25-MAR-1998; 98WO-US05956.
XX
PR 04-AUG-1997; 97US-0905359.
PR 25-MAR-1997; 97US-0041666.
PR 30-APR-1997; 97US-0045211.
PR 12-MAY-1997; 97US-0046256.
XX
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX
PI Affholter JA, Arnold FH, Giver LJ, Shao Z, Zhao H;
XX
DR WPI; 1998-542276/46.
XX
PT Producing double stranded mutagenised polynucleotides - by enzyme
PT catalysed DNA polymerisation from random or defined sequence
PT primers, useful for producing mutant genes with new or improved
XX
PS Example 7; Page 38; 68pp; English.
XX
CC This peptide is used in a novel method for producing double-stranded
CC mutagenised polynucleotides (PNS) from at least 1 template PN, where the
CC mutagenised PNS have at least 1 nucleotide which is different from the
CC same nucleotide in the template PN. The method is useful for the in vitro
CC mutagenesis and recombination of PNS based on polymerase-catalysed
CC extension of primers. The method can produce full-length genes which are
CC mutants of the original template PNS, and these genes can be further
CC amplified and cloned into vectors. The new PNS can be used in DNA-base
CC computing or for evolving new gene products with improved or new
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 24
AAW61357
ID AAW61357 standard; Protein; 4 AA.
XX
AC AAW61357;
XX
DT 25-SEP-1998 (first entry)
XX
DE Synthetic oligopeptide succinyl-AAPF-p-nitroanillide.
XX
KW alkaline protease; protease Q; denaturing agent resistance; detergent;
KW sodium dodecyl sulfate; urea.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site /note= "N-terminal succinyl"
FT Modified-site 4
FT Modified-site /note= "phosph-nitroanillidated"
XX
PN WO9821319-A1.
XX
PD 22-MAY-1998.
XX
PF 12-NOV-1997; 97WO-US21859.

XX 15-NOV-1996; 96US-0751070.
PR
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Damodaran S, Han X;
XX
DR WPI; 1998-322335/28.
XX
XX
PT New purified alkaline protease with strong proteolytic activity - is
PT highly resistant to denaturing agents and may be used in cleaning
PT compositions
XX
XX
PS Disclosure; Page 9; 50pp; English.
XX
XX The synthetic oligopeptide was used to determine the amidase activity of
CC protease Q. Protease Q is a pure alkaline protease composition which is
CC highly resistant to denaturing agents such as sodium dodecyl sulfate or
CC urea. It is therefore useful as an additive for detergent and cleaning
CC compositions, such as dish washing or laundry soap compositions. The
CC protease can also be used in industrial cleaning such as the cleaning of
CC ultrafiltration and microfiltration membranes in food and pharmaceutical
CC industries.
XX
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
|||
Db 1 AAPF 4

RESULT 25
AAW51451
ID AAW51451 standard; peptide; 4 AA.
XX
AC AAW51451;
XX
DT 02-SEP-1998 (first entry)
XX
DE Indicator for detection of leukocyte esterase activity in urine.
XX
XX Urine; Indicator; detection; leukocyte esterase; activity;
KW white blood cell.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Acetyl-, hydrochloride-H-, methoxysuccinyl-suc-,
FT suc-, glt-, mal- or HBrNH2(CH2)5CO-Ala"
FT Modified-site 4
FT /note= "Phe-nitroanilide"
XX
XX
PN US5776780-A.
XX
PD 07-JUL-1998.
XX
XX
PF 12-APR-1996; 96US-0631581.
XX
PR 12-APR-1996; 96US-0631581.
PR 28-MAY-1993; 93US-0068956.
PR 24-APR-1995; 95US-0429292.
XX
PA (CHIM-) CHIMERA RES & CHEM INC.
XX
PI Carter JM, Smith JV;
XX
DR WPI; 1998-398049/34.
XX

PT Automated measurement of white blood cell esterase activity in urine
PT - comprises adding reagent composition to the sample containing an
PT indicator, placing sample in a analyser and comparing absorbance
PT with a standard measurement
XX
XX
PS Claim 11; Column 14; 8pp; English.
XX
XX
CC The invention relates to a method for the measurement of white blood
CC cell esterase activity in urine. It comprises: (i) placing an aliquot of
CC the urine in an automated analyser sampling cup; (ii) placing the cup in
CC a sampling tray with the automated analyser, transferring the urine to a
CC cuvette and injecting at least one reagent composition in an aqueous
CC medium into the cuvette, where the reagent composition comprises a
CC buffer to adjust the pH of the urine to 7-9, an activator and surfactant
CC selected from benzalkonium chloride, 2,5-dimethylbenzenesulphonic acid
CC and benzethonium chloride, and an indicator to determine leukocyte
CC esterase activity in the urine; (iii) reading the aliquot of urine at
CC specified intervals, in accordance with a preprogrammed code introduced
CC into the automated analyser, at a preprogrammed monochromatically
CC specified wavelength, to compare absorbance of the urine sample and
CC reagent composition complex with that of a standard containing a known
CC concentration of leukocyte esterase and determining the quantitative
CC amount of leukocyte esterase in the patient's urine. The method provides
CC a fully automated method for determining white blood cells in urine. The
CC method is preferable to the old dipstick method as it reduces consumable
CC material and labour costs. It also offers increased accuracy, sensitivity
CC and reduction of interference by substances which affected prior art
CC tests. The present sequence represents a specifically claimed indicator.
XX
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
|||
Db 1 AAPF 4

RESULT 26
AAW51610
ID AAW51610 standard; peptide; 4 AA.
XX
AC AAW51610;
XX
DT 03-SEP-1998 (first entry)
XX
DE Peptide conjugated to lipid for use in liposomal drug delivery.
XX
XX Liposome; conjugate; drug delivery; peptidase-secreting cell;
KW tumour; diagnosis; therapy.
XX
OS Synthetic.
XX
XX
PN WO9816240-A1.
XX
PD 23-APR-1998.
XX
XX
PF 15-OCT-1997; 97WO-US18538.
XX
PR 15-OCT-1996; 96US-0027544.
XX
XX
PA (LIPO) LIPOSOME CO INC.
XX
PI Ali S, Cabralilly D, Erukulla RK, Franklin JC;
PI Janoff AS, Meers PR, Pak C;
XX
XX
DR WPI; 1998-261025/23.
XX
XX

New peptide-lipid conjugates are incorporated into liposome(s) - to
selectively destabilise the liposome(s) in the vicinity of target
peptidase-secreting cells, e.g., tumour cells; useful in diagnosis

PT and therapy
XX
PS Claim 6; Page 33; 55pp; English.
XX
CC The invention relates to peptide-lipid conjugates and their use in
CC the preparation of liposomes which are predisposed to degradation in the
CC presence of peptidase-secreting cells and hence are targeted to these
CC cells. The liposomes can be used to treat mammalian diseases, disorders
CC and conditions, e.g. tumours, microbial infections and inflammation. The
CC liposomes allow selective delivery of an active agent to desired cells.
CC The liposomes are stable when the peptide remains conjugated to the
CC lipid. However, once the peptide portion of the conjugate is cleaved
CC from the lipid, by the action of cell-secreted peptidases, the liposomes
CC destabilise and release their contents in the vicinity of, or into, the
CC secreting cells. The liposomes can thus be used to treat conditions
CC characterised by the occurrence of peptidase-secreting cells. For
CC example, the breast cancer cell line MCF-7 is known to secrete elastase,
CC the levels of which are inversely correlated to overall survival in
CC breast cancer patients. The present sequence represents a specifically
CC claimed peptide which forms part of the peptide-lipid conjugate.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
 ||||
Db 1 AAPF 4

RESULT 27
AAM48424
ID AAM48424 standard; Peptide; 4 AA.
XX
AC AAM48424;
XX
DT 16-JUL-2002 (first entry)
XX
DE Alkaline protease-related peptide #2.
XX
KM Alkaline protease; enzyme.
XX
OS Unidentified.
XX
PN KR9709083-B1.
XX
PD 05-JUN-1997.
XX
PF 28-APR-1994; 94KR-0009139.
XX
PR 28-APR-1994; 94KR-0009139.
XX
PA (KOAD) KAIST.
XX
PI Yu W, Lee H, Jang W, Kim E;
XX
DR WPI; 1999-523621/44.
XX
PT Novel alkaline protease gene and expression thereof in e.coli -
PT NoAbstract
XX
PS Disclosure; Page 4; 17pp; Korean.
XX
CC The present invention relates to a novel alkaline protease gene and a
CC method for expression thereof in E.coli. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
 ||||
Db 1 AAPF 4

RESULT 28
ABB07072
ID ABB07072 standard; peptide; 4 AA.
XX
AC ABB07072;
XX
DT 28-JUN-2002 (first entry)
XX
DE Alkaline protease related peptide.
XX
KW Alkaline protease; enzyme; microorganism.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT FT /note= "succinylated"
FT Modified-site 4 /note= "C-terminally attached to pNA"
XX
PN KR135392-B1.
XX
PD 23-APR-1998.
XX
PF 28-APR-1994; 94KR-0009140.
XX
PR 28-APR-1994; 94KR-0009140.
XX
PA (KOAD) KAIST.
XX
PI Yoo W, Lee H, Jang W, Kim E, Kang J;
XX
DR WPI; 1999-630649/54.
XX
PT NOVEL MICROORGANISM PRODUCING NOVEL ALKALINE PROTEASE -
XX
PS Example 3; Page 6; 9pp; Korean.
XX
CC The present invention describes a microorganism producing an alkaline
CC protease. The present sequence represents a peptide used in the
CC exemplification of the present invention.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
 ||||
Db 1 AAPF 4

RESULT 29
AAW84189
ID AAW84189 standard; peptide; 4 AA.
XX
AC AAW84189;
XX
DT 25-MAR-1999 (first entry)
XX
DE Peptide comprising a proteinase site.
XX
KW Proteinase site; bone morphogenetic fusion protein; bone binding site;
KW bone morphogenetic protein; transforming growth factor beta;
KW active fragment; wound healing; bone growth.

XX Unidentified.
XX
PN WO9855137-A1.
XX
PD 10-DEC-1998.
XX
PF 02-JUN-1998; 98WO-US11189.
XX
PR 03-JUN-1997; 97US-0868452.
XX
PA (HALL/) HALL F L.
PA (HANB/) HAN B.
PA (NIMN/) NIMNI M E.
PA (SHOR/) SHORS E C.
PA (WULL/) WU L.
XX
PI Hall FL, Han B, Nimni ME, Shors EC, Wu L;
XX
DR WPI; 1999-059875/05.
XX
PS Claim 8; Page 34; 64pp; English.
XX
PT New bone morphogenetic fusion proteins - comprising a purification
PT tag and a bone morphogenetic active fragment, used for enhancing
PT wound healing or bone growth
XX
CC The present peptide represents a proteinase site used in the creation
CC of the bone morphogenetic fusion proteins of the invention. The bone
CC morphogenetic fusion protein may contain some or all of the following
CC elements: a purification tag, a proteinase site, an ECM/bone binding
CC site, a second proteinase site, and a bone morphogenetic protein
CC active fragment. The fusion proteins of the invention also includes
CC proteins that have transforming growth factor beta active fragments
CC instead of bone morphogenetic protein active fragments. The bone
CC morphogenetic fusion proteins can be used for enhancing wound healing
CC or bone growth.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
 ||||
Db 1 AAPF 4

RESULT 30
AAB20774
ID AAB20774 standard; peptide; 4 AA.
XX
AC AAB20774;
XX
DT 21-DEC-2000 (first entry)
XX
DE Chymotrypsin enzyme substrate peptide sequence.
XX
KW Heteroaryl amidine; methylamidine; guanidine; protease inhibitor;
KW enzyme substrate; serine protease; chymotrypsin; trypsin; plasmin;
KW urokinase; cytosolic; antiproliferative; antiproliferative; antiproliferative;
KW neuroprotective; antiinflammatory; antirheumatic; antiarthritic;
KW antiarteriosclerotic; antiparkinsonian; vasotropic; restenosis;
KW benign prostatic hypertrophy; prostatic carcinoma; tumour metastasis;
KW psoriasis; adult respiratory distress syndrome; wound healing; gout;
KW rheumatoid arthritis; reperfusion damage; atherosclerosis; neoplasia;
KW metastasis; emphysema; Alzheimer's disease; pancreatitis;
KW Parkinson's disease.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers

FT Modified-site 1
FT /note= "succinylated"
FT Modified-site 4
FT /note= "Arg is C-terminally modified to
FT Arg-p-nitroanilide"
XX
PN WO200047578-A1.
XX
PD 17-AUG-2000.
XX
PF 11-AUG-1999; 99WO-US18065.
XX
PR 09-FEB-1999; 99US-0247062.
XX
PA (THRE-) 3-DIMENSIONAL PHARM INC.
XX
PI Illig CR, Subasinghe NL, Hoffman JB, Wilson KJ, Rudolph MJ;
PI Marugan JI;
XX
DR WPI; 2000-558186/51.
XX
PS Example 292; Page 277; 326pp; English.
XX
PT New heteroaryl amidines, methylamidines and guanidines, useful for
PT treating e.g. benign prostatic hypertrophy, adult respiratory distress
PT syndrome, wound healing, gout and rheumatoid arthritis
XX
CC The present invention describes heteroaryl amidines, methylamidines and
CC guanidines (I) and their solvates, hydrates and salts. (I) can have
CC cytosolic, antiproliferative, antiproliferative, neuroprotective,
CC antiinflammatory, vulnary, antirheumatic, antiarthritic,
CC antiarteriosclerotic, antiparkinsonian and vasotropic activities.
CC (I) are inhibitors of proteases selected from leukocyte neutrophil
CC elastase, chymotrypsin, trypsin, pancreatic elastase, cathepsin G,
CC thrombin, urokinase, factor Xa, plasmin, thermolysin, C-1 esterase,
CC C-3 convertase, acrosin, thrombin, kallikreins, and pepsin, especially
CC trypsin, chymotrypsin, plasmin or urokinase. (I) can be used for
CC treating benign prostatic hypertrophy, prostatic carcinoma, tumour
CC metastasis, restenosis or psoriasis; adult respiratory distress
CC syndrome, wound healing, gout, rheumatoid arthritis, reperfusion
CC damage, atherosclerosis, neoplasia, metastasis, emphysema,
CC Alzheimer's disease, pancreatitis, or Parkinson's disease. The present
CC sequence represents an enzyme substrate peptide which is used in an
CC example from the present invention for the in vitro inhibition of
CC purified enzymes.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 21; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
 ||||
Db 1 AAPF 4

RESULT 31
AAB03092
ID AAB03092 standard; peptide; 4 AA.
XX
AC AAB03092;
XX
DT 10-OCT-2000 (first entry)
XX
DE Substrate peptide #3.
XX
KW Serine protease; trypsin activity; Trichoderma; bacterial; coagulant;
KW hypertensive; antiinflammatory; leather preparation; silk treatment.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers

FT Modified-site 1 /note= "Conjugated to succinyl moiety"
FT Modified-site 4 /note= "Conjugated to p-nitroaniline (pNA)"
XX Jp2000116377-A.
XX 25-APR-2000.
XX
PF 08-OCT-1998; 98JP-0303263.
XX
PR 08-OCT-1998; 98JP-0303263.
XX
PA (AMAN) AMANO PHARM KK.
XX
DR WPI; 2000-369402/32.
XX
PT A new serine protease and its preparation, used clinically in blood
PT coagulation, hypotension and anti-inflammation -
XX
PS Example 3; Page 5; 9pp; Japanese.
XX
CC The invention relates to a novel serine protease from Trichoderma sp.
CC No.9064. The N-terminus of this protease is given in AAB03085. The novel
CC protease has trypsin-like activity, specifically cleaving peptide chains
CC on the carboxyl side of a basic amino acid (e.g., arginine or lysine).
CC It has a pH optimum of 7-8, and is stable between pH 5 and pH10. It has
CC a temperature optimum of approximately 40 degrees Celsius, and is stable
CC between 40 and 50 degrees Celsius. The invention also relates to a
CC method for the preparation of the novel serine protease, and the use of
CC the protease in protein degradation. The enzyme has coagulant,
CC hypertensive and anti-inflammatory effects. It may also be used in the
CC preparation of leather, for raw silk treatment and for the preparation of
CC protein hydrolysate. Sequences AAB03090-B03094 represent tetrapeptides
CC used in the determination of the activity of the novel protease.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 21; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
|||
Db 1 AAPF 4

RESULT 32
AAB01922
ID AAB01922 standard; peptide; 4 AA.
XX
AC AAB01922;
XX
DT 18-SEP-2000 (first entry)
XX
DE Synthetic PPIase substrate peptide.
XX
KW Peptidyl prolyl cis-trans isomerase; PPIase; hepatitis B surface antigen;
KW HBSAg; hepatitis B virus; HBV; granulation; immunoassay;
KW substrate peptide.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminus is conjugated to a succinyl moiety"
FT Modified-site 4 /note= "C-terminus is conjugated to a p-nitroanilide
FT moiety"
XX
PN Jp2000105234-A.
XX
PR 11-APR-2000.
PD

XX
PF 29-SEP-1998; 98JP-0275008.
XX
PR 29-SEP-1998; 98JP-0275008.
XX
PA (SEKI) SEKISUI CHEM IND CO LTD.
PA (KAIT-) KAITO BIOTECHNOLOGY KENKYUSHO KK.
XX
DR WPI; 2000-334401/29.
XX
PT Preparation of small granules of Hepatitis B surface antigen, used in
PT manufacture of immunoassay reagent, by simultaneous granulating and
PT cis-trans isomerase reaction of antigen -
XX
PS Disclosure; Page 5; 14pp; Japanese.
XX
CC The invention relates to the preparation of small granules of hepatitis
CC B surface antigen (HBSAg) by granulating the antigen with a surfactant,
CC reducer or protein denaturant. The granulated antigen is simultaneously
CC subjected to cis-trans prolyl isomerisation by contacting it with a
CC peptidyl prolyl cis-trans isomerase (PPIase). The small granulated HBSAg
CC is used in the manufacture of a hepatitis B virus (HBV) immunoassay
CC reagent. Small granules comprising activated HBSAg, an enzyme labelled
CC antigen reagent and a latex aggregation reagent can be obtained. These
CC can be used in a highly sensitive HBV immunoassay. The present
CC sequence represents a PPIase synthetic substrate peptide referred to in
CC the disclosure of the invention.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 21; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
|||
Db 1 AAPF 4

RESULT 33
AAV91025
ID AAV91025 standard; peptide; 4 AA.
XX
AC AAV91025;
XX
DT 06-SEP-2000 (first entry)
XX
DE Model substrate peptide sequence.
XX
KW Carboxylic acid; isostere; heterocyclic compound; vision disorder;
KW memory disorder; ophthalmological; nootropic; rotamase inhibitor;
KW peptidylprolyl isomerase inhibitor; visual impairment; orbital disorder;
KW lacrimal apparatus; eyelid; conjunctiva; cornea; cataract; uveal tract;
KW retina; optic nerve; visual pathway; free radical induced eye disorder;
KW immunologically-mediated eye disorder; ophthalmologic disorder.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "succinylated"
FT Modified-site 4 /note= "Phe is C-terminally modified to
FT Phe-p-nitroanilide"
XX
PN WO200009102-A2.
XX
PD 24-FEB-2000.
XX
PF 12-AUG-1999; 99WO-US18230.
XX
PR 14-AUG-1998; 98US-0134472.
PD

PA (GUIL-) GUILFORD PHARM INC.
XX
PI Ross DT, Sauer H, Hamilton GS, Steiner JP;
XX
DR WPI; 2000-224156/19.
XX
PT Treating a vision disorder or enhancing memory by administration of
PT N-heterocyclic ring compound including a direct or indirect carboxyl
PT substituent -
PS Disclosure; Page 37; 99pp; English.
XX
CC The present invention describes N-heterocyclic ring compounds containing
CC carboxylic acid or carboxylic acid isostere attached to 2-carbon of
CC N-heterocyclic ring useful for treating vision disorders or enhancing
CC memory performance. The compounds have ophthalmological and neurotropic
CC activities and are peptidylprolyl isomerase and rotamase inhibitors.
CC Methods from the present invention are useful for treating vision
CC disorders such as visual impairments, orbital disorders, disorders of
CC the lacrimal apparatus, disorders of the eyelids, disorders of the
CC conjunctiva, disorders of the cornea, cataract, disorders of the uveal
CC tract, disorders of the retina, disorders of the optic nerve or visual
CC pathways, free radical induced eye disorders and diseases,
CC immunologically-mediated eye disorders and disorders, eye injuries and
CC symptoms and complications of eye disease, eye disorder or eye injury.
CC The methods are also useful for improving naturally-occurring vision,
CC in the absence of any ophthalmologic disorder, disease or injury. The
CC compositions are additionally useful for treating memory impairment or
CC enhancing memory performance in an animal. The compounds may be
CC differentiated from the non-immunosuppressive compounds used to treat
CC vision disorders by their novel small molecule structure and their lack
CC of general, systemic effects. The present sequence represents a model
CC substrate peptide sequence which is used in the exemplification of the
CC present invention.
XX
SQ Sequence 4 AA;
OY Query Match 100.0%; Score 21; DB 21; Length 4;
Db Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AAPF 4
Db 1 AAPF 4
RESULT 34
AAY97813
ID AAY97813 standard; protein; 4 AA.
XX
AC AAY97813;
XX
DT 21-AUG-2000 (first entry)
XX
DE Protease peptide substrate.
XX
KW Pmrotease substrate; immunogenicity; allergenicity;
KW industrial product; pharmaceutical; hypoallergenic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "Conjugated to a succinyl moiety"
FT Modified-site 4 /note= "Conjugated to a nitro-anilide moiety"
XX
XX
PN WO200022103-A1.
XX
PD 20-APR-2000.
XX
PE 12-OCT-1999; 99WO-DK00542.
XX

PR 13-OCT-1998; 98DK-0001301.
PR 04-OCT-1999; 99DK-0001418.
XX
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Olsen AA, von der Osten C, Andersen KV, Ernst S, Roggen EL;
XX
DR WPI; 2000-329163/28.
XX
PT New modified polypeptides having an attached polymer for reducing
PT immune responses, useful in e.g. detergents, cleaning products, skin
PT care products, food or feed products, textile products or
PT pharmaceuticals -
XX
PS Example 5; Page 49; 108pp; English.
XX
CC The invention relates to proteins with reduced immunogenicity, having
CC one or more modified amino acids, where the alpha-carbon atoms of the
CC amino acids are located less than 15 Angstroms from a ligand bound to
CC the protein. The modification may entail substitution of the endogenous
CC residue for a non-endogenous residue and/or attachment of polymeric
CC molecules such as carbohydrates or branched polyethylene glycols to the
CC amino acid residues. The residue to be modified is identified from the
CC 3-dimensional structure of the protein determined, for example, by X-ray
CC crystallography or NMR. A wide variety of enzymes may be modified
CC according to the invention, including proteases (especially subtilisins),
CC carboxydases (such as amylase), isomerases, transferases and
CC oxidoreductases. The modified proteins of the invention may be used
CC for reducing the allergenicity of industrial products (i.e., those which
CC are not intended to enter the circulatory system. The proteins may be
CC used in cleaning agents (such as laundry products, dish-washing products
CC or hard surface cleaning products), skin-care products, textile treatment
CC products (e.g., bleaching agents) and food products. The modified
CC polypeptides can also be used for reducing the immunogenicity of
CC pharmaceuticals. The modified proteins have reduced immunogenicity or
CC allergenicity while maintaining a high percentage of activity. The
CC present sequence represents a synthetic peptide substrate used in
CC assays to determine the activities of modified proteases of the
CC invention.
XX
SQ Sequence 4 AA;
OY Query Match 100.0%; Score 21; DB 21; Length 4;
Db Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AAPF 4
Db 1 AAPF 4
RESULT 35
AAY80465
ID AAY80465 standard; peptide; 4 AA.
XX
AC AAY80465;
XX
DT 06-JUN-2000 (first entry)
XX
DE Chymase inhibitor peptide.
XX
KW Chymase inhibitor; N-substituted azetidinone compound; inflammation;
KW allergy; circulation disorder; restenosis; fibrosis; rheumatism; asthma;
KW dermatitis; arthritis; psoriasis; hepatitis; sclerosis; conjunctivitis;
KW Crohn's disease; hypertension; myocardial disease; heart failure;
KW kidney damage; stroke; arteriosclerosis; immunomodulator.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "Suc-Ala"
FT Modified-site 4
XX

FT /note= "Phe-pNA"

XX WO200005204-A1.

PN 03-FEB-2000.

XX 16-JUL-1999; 99WO-JP03864.

XX 23-JUL-1998; 98JP-0207540.

XX (SHIO) SHIONOGI & CO LTD.

XX Uenaka M, Kii M, Nakajima M;

XX WPI; 2000-195090/17.

XX Chymase inhibitors containing N-substituted azetidinone compounds,

PT useful for treating, e.g. inflammation and asthma -

XX Disclosure; Page 150; 173pp; Japanese.

XX The invention relates to new chymase inhibitors containing an

CC N-substituted azetidinone compound or derivative. This sequence

CC represents an example of such an inhibitor. The inhibitors are useful

CC for treating inflammation, allergies, and circulation disorders,

CC including scarring, and keloid formation, post-myocardial infarct

CC chronic inflammation, post-surgical restenosis, fibrosis, rheumatism,

CC asthma, dermatitis, arthritis, psoriasis, hepatitis, hepatic sclerosis,

CC conjunctivitis, Crohn's disease, hypertension, myocardial disease,

CC heart failure, peripheral circulation disease, diabetic and non-diabetic

CC kidney damage, stroke, arteriosclerosis. The inhibitors may also be

CC used as immunomodulators.

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 21; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4

Db 1 AAPF 4

RESULT 36

AAV76808

ID AAV76808 standard; peptide; 4 AA.

XX AAV76808;

AC 26-APR-2000 (first entry)

XX Chymotrypsin substrate peptide.

DE Proteolytic enzyme; inhibitor; substrate: aminoguanidine; pyrazinone;

XX KW hydrazinoamide; alkoxyguanidine; alkoxyaminoamide; viral infection;

KW thrombotic disease; ischaemia; stroke; cancer; restenosis; septic shock;

KW myocardial infarction; disseminated intramuscular coagulopathy; sepsis;

KW unstable angina; coronary artery bypass; hip replacement; haemodialysis;

KW adult respiratory distress syndrome; rheumatoid arthritis; induration;

KW ulcerative colitis; metastasis; hypercoagulability; Alzheimer's disease;

KW Down's syndrome; fibrin formation; wound healing; inflammation; therapy;

KW Chymotrypsin.

XX Synthetic.

OS Key 1 Location/Qualifiers

XX Modified-site 1 /note= "N-succinyl-Ala"

FT Modified-site 4 /note= "Phe-p-nitroanilide"

XX WO964446-A1.

XX 16-DEC-1999.

XX 11-JUN-1999; 99WO-US13228.

XX 11-JUN-1998; 98US-0088989.

XX (THRE-) 3-DIMENSIONAL PHARM INC.

PA (LUTT/) LU T.

PA (TOMC/) TOMCZUK B E.

PA (MARK/) MARKOTAN T P.

XX Lu T, Tomczuk BE, Markotan TP;

XX WPI; 2000-147091/13.

XX Pyrazinone compounds used as potent thrombin (protease) inhibitors -

PT Example 36; Page 78; 112pp; English.

XX This sequence represents a peptide substrate for chymotrypsin.

CC The invention relates to aminoguanidine (hydrazinoamide) and

CC alkoxyguanidine (alkoxyaminoamide) pyrazinones that function as

CC proteolytic enzyme inhibitors. A pharmaceutical composition containing

CC the pyrazinone is used to treat thrombotic diseases associated with

CC ischaemia, viral infections, stroke, cancer, restenosis, myocardial

CC infarction, disseminated intramuscular coagulopathy which occurs during

CC septic shock, unstable angina, disseminated intramuscular coagulation

CC caused by trauma, coronary artery bypass, hip replacement, thrombolytic

CC therapy, sepsis, haemodialysis, adult respiratory distress syndrome,

CC rheumatoid arthritis, ulcerative colitis, induration, metastasis,

CC hypercoagulability during chemotherapy (Alzheimer's disease, Down's

CC syndrome, fibrin formation in the eye, wound healing, or inflammation in

CC a mammal.

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 21; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4

Db 1 AAPF 4

RESULT 37

AAV78789

ID AAV78789 standard; peptide; 4 AA.

XX AAV78789;

AC 09-MAY-2000 (first entry)

XX Subtilisin substrate peptide.

DE Subtilisin substrate peptide.

XX Subtilisin substrate; chemically modified mutant protein; anti-microbial;

KW reduced amidase activity; chiral resolution; regioselective acetylation;

KW pharmaceutical; diagnostic agent; cleaning composition; brewing;

KW textile treatment.

XX Synthetic.

OS Key 1 Location/Qualifiers

XX Modified-site 1 /note= "Succinyl-Ala"

FT Modified-site 4 /note= "Phe-p-nitroanilide"

XX WO200001712-A2.

XX 13-JAN-2000.

PF 02-JUL-1999; 99WO-US15138.
 XX
 PR 02-JUL-1998; 98US-0091687.
 PR 28-APR-1999; 99US-0131446.
 XX
 PA (GEMV) GENENCOR INT INC.
 XX
 PI Jones JB, Davis BG;
 XX
 DR WPI; 2000-170996/15.
 XX
 PT New chemically modified mutant proteins, particularly proteases, for
 PT e.g. peptide synthesis -
 XX
 PS Example 2; Page 33; 86pp; English.
 XX

CC This sequence represents a subtilisin peptide substrate. The peptide is
 CC used in a procedure for the modification of subtilisin Bacillus lentus
 CC mutants. Subtilisin mutants are used in the invention which relates to a
 CC chemically modified mutant protein which includes a Cys residue which
 CC replaces another amino acid in a precursor, that has been subsequently
 CC modified by reaction with a glycosylated thiosulfonate. The chemically
 CC modified mutant proteins have altered functional properties,
 CC e.g. solubility; cell-cell signaling; catalytic, biological or
 CC pharmacological activities, also sites critical for immunological or
 CC allergic responses, or proteolytic degradation, can be masked. They may
 CC be used to determine structure-function relationships, also e.g.:
 CC (a) to alter the catalytic activity of enzymes;
 CC (b) to improve suitability for use in vaccines;
 CC (c) to reduce allergenicity;
 CC (d) to improve solubility (e.g. to facilitate recovery or formulation),
 CC or
 CC (e) to improve stability against proteolysis.
 CC A particular application is to modify subtilisins so that they have
 CC reduced amidase and increased esterase activities, making them useful in
 CC chiral resolution; regioselective acylation; and synthesis of peptides or
 CC glycoproteins, but more generally modified proteins of the invention may
 CC also be used as pharmaceutical or diagnostic agents, or in cleaning
 CC compositions, textile treatment, modification of foods or feeds, brewing
 CC or starch processing, as anti-microbials and in personal care
 CC formulations.
 CC
 XX

SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 21; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
 ||||
 Db 1 AAPF 4

RESULT 38

AAV49452
 ID AAV49452 standard; peptide; 4 AA.
 XX

AC AAV49452;

DT 17-MAR-2000 (first entry)

DE Cathepsin G peptide fragment.

KW Proteinase inhibitor; disulfide core proteinase inhibitor; Zdscl; mouse;
 KW pancreatitis; shock syndrome; hyperfibrinolytic hemorrhage; gene therapy;
 KW myocardial infarction; antinflammatory; vasotropic; cardiant;
 KW Cathepsin G.
 XX

OS Homo sapiens.

PN WO9963091-A1.

PD 09-DEC-1999.

XX
 PF 04-JUN-1999; 99WO-US12545.
 XX
 PR 04-JUN-1998; 98US-0090895.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI ConKlin DC;
 XX
 DR WPI; 2000-072884/06.
 XX
 PT Novel polypeptide sequences used as proteinase inhibitors -
 PT
 XX
 PS Disclosure; Page 13; 85pp; English.
 XX

CC The invention relates to a new class of proteinase inhibitors called
 CC disulfide core proteinase inhibitors (Zdscl) and provides murine and
 CC human Zdscl polypeptides. The disulfide core proteins are proteinase
 CC inhibitors used for the treatment of acute pancreatitis, various stages
 CC of shock syndrome, hyperfibrinolytic hemorrhage and myocardial
 CC infarction. Antagonists of the polypeptides can be used as research
 CC reagents for characterizing sites of ligand-receptor interaction.
 CC Antibodies against the proteins may be used for tagging cells that
 CC express the protein, for detecting the protein, and for screening
 CC expression libraries. Polynucleotides encoding the polypeptides can be
 CC used in gene therapy applications where it is desired to increase or
 CC inhibit the protein activity.
 CC
 XX

SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 21; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
 ||||
 Db 1 AAPF 4

RESULT 39

AAV59632
 ID AAV59632 standard; peptide; 4 AA.
 XX

AC AAV59632;

DT 23-MAR-2000 (first entry)

DE Peptidylprolyl cis-trans isomerase (PPIase) peptide.

KW Peptidylprolyl cis-trans isomerase; PPIase; protein stabilising;
 KW protein inactivation inhibitor.
 XX

OS Methanococcus voltae.

FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-succinyl-Ala"

FT Modified-site 4 /note= "phe-p-nitroanilide"

PN JP11302297-A.

PD 02-NOV-1999.

PF 16-APR-1998; 98JP-0105953.

PR 16-APR-1998; 98JP-0105953.

PA (SEKI) SEKISUI CHEM IND CO LTD.

PA (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.

DR WPI; 2000-075343/07.

PT A method for stabilisation of a protein - using peptidylprolyl
PT cis-trans isomerase from a Methanococcus genus microbe
XX
PS Disclosure; Page 3; 7pp; Japanese.
XX
CC This sequence is a peptide fragment of Methanococcus voltae
CC peptidylprolyl cis-trans isomerase (PPIase). The peptide is used in the
CC invention which relates to a method for the stabilisation of a protein.
CC In the method, the PPIase is prepared from a Methanococcus genus microbe
CC or is expressed from DNA prepared from a Methanococcus genus microbe. The
CC PPIase is present in a solution that is used to inhibit inactivation of a
CC protein present in the solution. The method is useful for stabilising a
CC protein. The method is versatile and does not require the use of a high
CC energy substance such as ATP.
XX
SQ Sequence 4 AA;
QY 1 AAPF 4
Db 1 AAPF 4
Query Match 100.0%; Score 21; DB 21; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 40
AAU07699
ID AAU07699 standard; Peptide; 4 AA.
XX
AC AAU07699;
XX
DT 18-DEC-2001 (first entry)
XX
DE Ztryp3 substrate used in serine protease activity assay.
XX
KW Human; Ztryp3; serine protease; asthma; vascular function; inflammation;
KW gene therapy; stroke; testicular function; spermatogenesis; haemostatic;
KW mass spectrometry; circular dichroism; X-ray crystallography;
KW nuclear magnetic resonance spectroscopy; antiasthmatic; antiinflammatory;
KW cerebroprotective.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "Suc-Ala"
FT Modified-site 4 /note= "Phe-pNA"
FT
XX
PN WO200166771-A2.
XX
PD 13-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06432.
XX
PR 03-MAR-2000; 2000US-0518387.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Conklin DC;
PI
DR WPI; 2001-589946/66.
XX
XX
PT Novel Ztryp3 polypeptides and polynucleotides useful in the treatment
PT of asthma, vascular disorders including stroke, inflammation and
PT testicular function
XX
PS Disclosure; Page 39; 82pp; English.
XX
CC The invention relates to an isolated human Ztryp3 polypeptide, a member
CC of the serine protease family. Ztryp3 polypeptides and their associated

CC polynucleotides are useful in diagnosis, therapy and industry and are
CC used as targets for identifying modulators, preferably inhibitors of
CC serine protease activity. The sequences are useful in the treatment of
CC asthma, vascular function such as stroke, inflammation and testicular
CC function (by modulating spermatogenesis). Ztryp3 proteins can be used for
CC identifying peptide cleavage sites and for coupling amino and carboxy
CC terminal tags. The polypeptides are also useful to teach analytical
CC skills such as mass spectrometry, circular dichroism, X-ray
CC crystallography and nuclear magnetic resonance spectroscopy. This
CC sequence represents a Ztryp3 substrate used in serine protease activity
CC assays.
XX
SQ Sequence 4 AA;
QY 1 AAPF 4
Db 1 AAPF 4
Query Match 100.0%; Score 21; DB 22; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: December 6, 2002, 13:28:01
Job time : 26 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:26:14 ; Search time 9.33333 Seconds
(without alignments)
12.610 Million cell updates/sec

Title: US-10-033-526-1
Perfect score: 21
Sequence: 1 AAPF 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match length | DB ID | Description |
|------------|-------|--------------------|-------|-------------------------------------|
| 1 | 21 | 100.0 | 4 1 | US-07-973-335-1 Sequence 1, Appl |
| 2 | 21 | 100.0 | 4 1 | US-08-153-331-19 Sequence 19, Appl |
| 3 | 21 | 100.0 | 4 1 | US-08-004-643C-2 Sequence 2, Appl |
| 4 | 21 | 100.0 | 4 1 | US-08-462-456-5 Sequence 5, Appl |
| 5 | 21 | 100.0 | 4 1 | US-08-282-860-1 Sequence 1, Appl |
| 6 | 21 | 100.0 | 4 1 | US-08-407-000-6 Sequence 6, Appl |
| 7 | 21 | 100.0 | 4 1 | US-07-890-422B-28 Sequence 28, Appl |
| 8 | 21 | 100.0 | 4 1 | US-08-276-936A-2 Sequence 2, Appl |
| 9 | 21 | 100.0 | 4 1 | US-08-345-820B-2 Sequence 3, Appl |
| 10 | 21 | 100.0 | 4 1 | US-08-544-143A-3 Sequence 3, Appl |
| 11 | 21 | 100.0 | 4 1 | US-08-397-602A-9 Sequence 9, Appl |
| 12 | 21 | 100.0 | 4 1 | US-08-437-029-5 Sequence 5, Appl |
| 13 | 21 | 100.0 | 4 1 | US-08-424-022-19 Sequence 19, Appl |
| 14 | 21 | 100.0 | 4 1 | US-08-439-534-10 Sequence 10, Appl |
| 15 | 21 | 100.0 | 4 1 | US-08-434-959-3 Sequence 3, Appl |
| 16 | 21 | 100.0 | 4 1 | US-08-434-959-4 Sequence 4, Appl |
| 17 | 21 | 100.0 | 4 1 | US-08-434-959-5 Sequence 5, Appl |
| 18 | 21 | 100.0 | 4 1 | US-08-434-959-6 Sequence 6, Appl |
| 19 | 21 | 100.0 | 4 1 | US-08-460-343B-56 Sequence 56, Appl |
| 20 | 21 | 100.0 | 4 1 | US-08-067-180B-2 Sequence 2, Appl |
| 21 | 21 | 100.0 | 4 1 | US-08-398-028B-56 Sequence 56, Appl |
| 22 | 21 | 100.0 | 4 1 | US-08-470-837-3 Sequence 3, Appl |
| 23 | 21 | 100.0 | 4 2 | US-08-722-268-2 Sequence 2, Appl |
| 24 | 21 | 100.0 | 4 2 | US-08-504-265B-56 Sequence 56, Appl |
| 25 | 21 | 100.0 | 4 2 | US-08-439-005-10 Sequence 10, Appl |
| 26 | 21 | 100.0 | 4 2 | US-08-698-575E-2 Sequence 2, Appl |
| 27 | 21 | 100.0 | 4 2 | US-08-424-017B-19 Sequence 19, Appl |

| | | | | |
|----|----|-------|-----|-------------------------------------|
| 28 | 21 | 100.0 | 4 2 | US-08-751-070B-4 Sequence 4, Appl |
| 29 | 21 | 100.0 | 4 3 | US-09-069-823-2 Sequence 2, Appl |
| 30 | 21 | 100.0 | 4 3 | US-09-076-460-1 Sequence 1, Appl |
| 31 | 21 | 100.0 | 4 3 | US-08-950-618-3 Sequence 3, Appl |
| 32 | 21 | 100.0 | 4 3 | US-08-950-618-18 Sequence 18, Appl |
| 33 | 21 | 100.0 | 4 3 | US-08-950-618-31 Sequence 31, Appl |
| 34 | 21 | 100.0 | 4 3 | US-09-120-365-101 Sequence 101, App |
| 35 | 21 | 100.0 | 4 3 | US-08-319-501-10 Sequence 10, Appl |
| 36 | 21 | 100.0 | 4 4 | US-09-168-010-3 Sequence 3, Appl |
| 37 | 21 | 100.0 | 4 4 | US-09-168-010-18 Sequence 18, Appl |
| 38 | 21 | 100.0 | 4 4 | US-09-168-010-31 Sequence 31, Appl |
| 39 | 21 | 100.0 | 4 4 | US-08-905-359A-25 Sequence 25, Appl |
| 40 | 21 | 100.0 | 4 4 | US-09-353-556-25 Sequence 25, Appl |
| 41 | 21 | 100.0 | 4 4 | US-09-515-039-101 Sequence 101, App |
| 42 | 21 | 100.0 | 4 4 | US-09-147-502A-1 Sequence 1, Appl |
| 43 | 21 | 100.0 | 4 4 | US-09-326-039-15 Sequence 15, Appl |
| 44 | 21 | 100.0 | 4 4 | US-09-343-650-3 Sequence 3, Appl |
| 45 | 21 | 100.0 | 4 4 | US-09-343-650-18 Sequence 18, Appl |

. ALIGNMENTS

RESULT 1
US-07-973-335-1
; Sequence 1, Application US/07973335
; Patent No. 5338547
; GENERAL INFORMATION:
; APPLICANT: Kennedy and Szuhaj
; TITLE OF INVENTION: No. 5338547el Bowman-Birk Inhibitor
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz
; ADDRESSEE: Mackiewicz & No. 5338547ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/973,335
; FILING DATE: 19921102
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 824,719
; FILING DATE: January 17, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 579,155
; FILING DATE: September 6, 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: UPSC-0003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-07-973-335-1

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 1 AAPF 4

RESULT 2

US-08-155-331-19
; Sequence 19, Application US/08155331
; Patent No. 5441931
; GENERAL INFORMATION:
; APPLICANT: Foster, Donald C
; APPLICANT: Sprecher, Cindy
; APPLICANT: No. 5441931 is, Kjeld
; TITLE OF INVENTION: NOVEL HUMAN AMYLOID PROTEIN PRECURSOR
; TITLE OF INVENTION: HOMOLOG AND KUNITZ-TYPE INHIBITOR
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/155,331
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,692
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-684
; REFERENCE/DOCKET NUMBER: 92-21C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..4
; OTHER INFORMATION: /label= ALA-1
; OTHER INFORMATION: /note= "Amino terminal alanine residue is capped
; OTHER INFORMATION: with a succinyl group"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..4
; OTHER INFORMATION: /label= Phe-4
; OTHER INFORMATION: /note= "Carboxyl-terminal phenylalanine residue is
; OTHER INFORMATION: capped with p-nitroanil..."
US-08-155-331-19

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 1 AAPF 4

RESULT 3

US-08-004-643C-2

; Sequence 2, Application US/08004643C
; Patent No. 5480779
; GENERAL INFORMATION:
; APPLICANT: Gunter Fischer & Gerhard K Ilerz
; TITLE OF INVENTION: Cyclosporine Assay
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gunter Fischer et al. c/o G. P. Katona
; STREET: 230 Park Avenue, Room 2200
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10169
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: conf. to Patentin Release #1.0, Ver.#1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/004,643C
; FILING DATE: 12 January 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 703,590
; FILING DATE: 20 May 1991
; APPLICATION NUMBER: 398,092
; FILING DATE: 24 August 1989
; APPLICATION NUMBER: DD WP 601 F/319 577W
; FILING DATE: 07 September 1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Katona, Gabriel P.
; REGISTRATION NUMBER: 20,829
; REFERENCE/DOCKET NUMBER: 691-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-3377
; TELEFAX: (212)986-6126
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-004-643C-2

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 1 AAPF 4

RESULT 4

US-08-462-456-5
; Sequence 5, Application US/08462456
; Patent No. 5510333
; GENERAL INFORMATION:
; APPLICANT: Angelastro, Michael R
; APPLICANT: Bey, Philippe
; APPLICANT: Doherty, Niall S
; APPLICANT: Janusz, Michael J
; APPLICANT: Mehdi, Shujaath
; APPLICANT: Peet, No. 5510333 ton P
; TITLE OF INVENTION: Inhibitors of Cathepsin G and Elastase
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marion Merrell Dow Inc.
; STREET: 2110 East Galbraith Rd.
; CITY: Cincinnati P. O. Box 156300

```
STATE: Ohio
COUNTRY: USA
ZIP: 45215-6300
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,456
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/342,999
FILING DATE:
APPLICATION NUMBER: US/08/222,552
FILING DATE:
APPLICATION NUMBER: US/07/987,587
FILING DATE:
APPLICATION NUMBER: US/07/704,499
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Nesbitt, Stephen L
REGISTRATION NUMBER: 28,981
REFERENCE/DOCKET NUMBER: M01593
TELECOMMUNICATION INFORMATION:
TELEPHONE: (513) 948-7965
TELEFAX: (513) 948-7961
TELEX: 214320
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-462-456-5

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 4;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAFP 4
Db 1 AAFP 4

RESULT 5
US-08-282-860-1
Sequence 1, Application US/08282860
Patent No. 5561108
GENERAL INFORMATION:
APPLICANT: Tsay, Grace C.
APPLICANT: Cheung, Neal K. H.
TITLE OF INVENTION: Preparation of Alpha-1 Antichymotrypsin
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bayer Corporation
STREET: 800 Dwight Way
STREET: P. O. Box 1986
CITY: Berkeley
STATE: California
COUNTRY: USA
ZIP: 94701-1986
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb Storage
COMPUTER: IBM
OPERATING SYSTEM: DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/282,860
FILING DATE: 07/29/94
CLASSIFICATION: 514
```

```
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Giblin, James A.
REGISTRATION NUMBER: 25772
REFERENCE/DOCKET NUMBER: MSB-7214
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510)705-7910
TELEFAX: (510)705-7904
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4
TYPE: amino acid
STRANDEDNESS: single strand
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: peptide
FEATURE:
NAME/KEY: Modified peptide used for enzyme activity assay
OTHER INFORMATION: Peptide is modified with n-terminal
PUBLICATION INFORMATION:
AUTHORS: Largman, C.
AUTHORS: Broderick, J. W.
TITLE: A Sensitive New Substrate for Chymotrypsin
JOURNAL: Analytical Biochemistry
VOLUME: 99
PAGES: 316-329
DATE: 01-NOV-1979
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 4
US-08-282-860-1

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 4;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAFP 4
Db 1 AAFP 4

RESULT 6
US-08-407-000-6
Sequence 6, Application US/08407000
Patent No. 5578324
GENERAL INFORMATION:
APPLICANT: Dohi, Masahiko
APPLICANT: Nishibe, Yoshihisa
APPLICANT: Makino, Yuji
TITLE OF INVENTION: PEPTIDE PROTEINACEOUS DRUG NASAL
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/407,000
FILING DATE: 29-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/01257
FILING DATE: 29-JUL-1994
APPLICATION NUMBER: JP-A-5-206922
```

;; FILING DATE: 30-JUL-1993
;; APPLICATION NUMBER: JP-A-5-235841
;; FILING DATE: 30-AUG-1993
;; APPLICATION NUMBER: JP-A-6-1644
;; FILING DATE: 12-JAN-1994
;; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-407-000-6

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 7
US-07-890-422B-28
; Sequence 28, Application US/07890422B
; Patent No. 5602102

GENERAL INFORMATION:

;; APPLICANT: THIELE, DWAIN L.
;; APPLICANT: LIPSKY, PETER E.
;; APPLICANT: MCGUIRE, MICHAEL J.
;; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASE-I
;; TITLE OF INVENTION: INHIBITORS AND USES THEREOF
;; NUMBER OF SEQUENCES: 38
;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: ARNOLD, WHITE & DURKEE
;; STREET: P.O. BOX 4433
;; CITY: HOUSTON
;; STATE: TEXAS
;; COUNTRY: USA
;; ZIP: 77210

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: FLOPPY DISK
;; COMPUTER: IBM PC COMPATIBLE
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WORDPERFECT 5.1

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/890,422B
;; FILING DATE: 19920529

;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:

;; NAME: MAYFIELD, DENISE L.
;; REGISTRATION NUMBER: 33,732
;; REFERENCE/DOCKET NUMBER: UTSD:296/MAY
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 512-320-7200

;; TELEFAX: 512-474-7577

;; TELEX: NOT APPLICABLE

;; INFORMATION FOR SEQ ID NO: 28:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acid residues
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-890-422B-28

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 8
US-08-276-936A-2
; Sequence 2, Application US/08276936A
; Patent No. 5612194

GENERAL INFORMATION:

;; APPLICANT: Harvey Rubin, Barry Cooperman, No. 5612194man Schechter,
;; APPLICANT: Michael Plotkin, Zhi Wang
;; TITLE OF INVENTION: Methods of Producing Effective
;; TITLE OF INVENTION: Recombinant Serine Protease Inhibitors and Uses of These
;; NUMBER OF SEQUENCES: 30
;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Jane Massey Licata, Esq.
;; STREET: 210 Lake Drive East, Suite 201
;; CITY: Cherry Hill
;; STATE: NJ
;; COUNTRY: USA
;; ZIP: 08002

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE

;; COMPUTER: IBM 486

;; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS

;; SOFTWARE: WORDPERFECT 5.1

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/276,936A

;; FILING DATE: July 19, 1994

;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: 221,078

;; FILING DATE: March 31, 1994

;; APPLICATION NUMBER: 221,171

;; FILING DATE: March 31, 1994

;; APPLICATION NUMBER: 005,908

;; FILING DATE: January 15, 1993

;; APPLICATION NUMBER: 735,335

;; FILING DATE: July 24, 1991

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Jane Massey Licata

;; REGISTRATION NUMBER: 32,257

;; REFERENCE/DOCKET NUMBER: PENN-0027

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (609) 779-2400

;; TELEFAX: (609) 779-8488

;; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4

; TYPE: Amino Acid

; TOPOLOGY: linear
US-08-276-936A-2

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 9
US-08-345-820B-2

; Sequence 2, Application US/08345820B
; Patent No. 5618792

GENERAL INFORMATION:

;; APPLICANT: SUBSTITUTED HETEROCYCLIC COMPOUNDS USEFUL AS
;; TITLE OF INVENTION: INHIBITORS OF (SERINE PROTEASES) HUMAN NEUTROPHIL ELASTASE
;; NUMBER OF SEQUENCES: 4

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/345,820B
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 4 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : US-08-345-820B-2

Query Match 100.0%; Score 21; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
 Db 1 AAPF 4

RESULT 10
 US-08-544-143A-3

Sequence 3, Application US/08544143A

Patent No. 5646028

GENERAL INFORMATION:

APPLICANT: Leigh, Scott D.
 TITLE OF INVENTION: NOVEL ALKALINE PROTEASE AND USES
 THEREOF
 NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Majestic, Parsons, Siebert & Hsue
 STREET: Four Embarcadero Center, Suite 1100
 CITY: San Francisco
 STATE: California

COUNTRY: U.S.A.
 ZIP: 94111-4121

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/544,143A
 FILING DATE: 17-OCT-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: de Runtz, K. Alison
 REGISTRATION NUMBER: 37,119

REFERENCE/DOCKET NUMBER: 0409.054US3

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-362-5556
 TELEFAX: 415-362-5418

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
 LENGTH: 4 amino acids
 TYPE: amino acid
 STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: peptide
 : US-08-544-143A-3

Query Match 100.0%; Score 21; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
 Db 1 AAPF 4

RESULT 11
 US-08-397-602A-9

Sequence 9, Application US/08397602A

Patent No. 5646044

GENERAL INFORMATION:

APPLICANT: Wilson, Charles R
 APPLICANT: Tang, Maria R
 APPLICANT: Berger, Harald
 APPLICANT: Christianson, Teresa M

APPLICANT: Hansen, Dieter

TITLE OF INVENTION: Expression Systems for the Production
 of Target Proteins in Bacillus

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Henkel Corporation Law Department
 STREET: 140 Germantown Pike, Suite 150
 CITY: Plymouth Meeting
 STATE: PA

COUNTRY: USA
 ZIP: 19462

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/397,602A
 FILING DATE: 02-MAR-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jaeschke, Wayne C
 REGISTRATION NUMBER: 21,062
 REFERENCE/DOCKET NUMBER: D8969/M4828

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 832-2200

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
 LENGTH: 4 amino acids
 TYPE: amino acid
 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1

OTHER INFORMATION: /note= "N-SUCCINYL"

FEATURE: NAME/KEY: Modified-site

LOCATION: 4

OTHER INFORMATION: /note= "P-NITROANILIDE"

US-08-397-602A-9

Query Match 100.0%; Score 21; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
 Db 1 AAPF 4

RESULT 12
 US-08-437-029-5

Sequence 5, Application US/08437029

Patent No. 5668107

GENERAL INFORMATION:

APPLICANT: Miller, Edward J.

TITLE OF INVENTION: Compositions and Methods for
 Inhibiting Elastase

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew
 STREET: One Market Plaza, Stewart Tower, Suite 2000
 CITY: San Francisco

STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/437,029
FILING DATE: 08 May 95
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 017066-000620
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-437-029-5

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 1 AAPF 4

RESULT 13

US-08-424-022-19

Sequence 19, Application US/08424022

Patent No. 5677146

GENERAL INFORMATION:

APPLICANT: Foster, Donald C

APPLICANT: Sprecher, Cindy

APPLICANT: No. 5677146ris, Kjeld

TITLE OF INVENTION: NOVEL HUMAN AMYLOID PROTEIN PRECURSOR

TITLE OF INVENTION: HOMOLOG AND KUNITZ-TYPE INHIBITOR

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Zymogenetics, Inc.

STREET: 4225 Roosevelt Way, N.E.

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/424,022

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/985,692

FILING DATE: 02-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Parker, Gary E

REGISTRATION NUMBER: 31-684

REFERENCE/DOCKET NUMBER: 92-21C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-547-8080 ext 322

TELEFAX: 206-548-2329
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1..4
OTHER INFORMATION: /label= ALA-1
OTHER INFORMATION: /note= "Amino terminal alanine residue is capped
OTHER INFORMATION: with a succinyl group"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1..4
OTHER INFORMATION: /label= Phe-4
OTHER INFORMATION: /note= "Carboxyl-terminal phenylalanine residue is
OTHER INFORMATION: capped with p-nitroanil..."
US-08-424-022-19

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 1 AAPF 4

RESULT 14

US-08-439-534-10

Sequence 10, Application US/08439534

Patent No. 5719041

GENERAL INFORMATION:

APPLICANT: Lazarus, Robert A.

APPLICANT: Dennis, Mark S.

APPLICANT: Ulmer, Jana S.

TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING

TITLE OF INVENTION: ECOTIN AND HOMOLOGS THEREOF

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/439,534

FILING DATE: 11-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/319501

FILING DATE: 04-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/121004

FILING DATE: 14-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Rubinec, Jeffrey S.

REGISTRATION NUMBER: 36,575

REFERENCE/DOCKET NUMBER: P0859C1D2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-8228

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
 LENGTH: 4 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-439-534-10

Query Match 100.0%; Score 21; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
 ||||
 Db 1 AAPF 4

RESULT 15
 US-08-434-959-3
 ; Sequence .3, Application US/08434959
 ; Patent No. 5736520
 ; GENERAL INFORMATION:
 ; APPLICANT: Bey, Philippe
 ; APPLICANT: Angelastro, Michael R
 ; APPLICANT: Mehdi, Shujaath
 ; TITLE OF INVENTION: No. 5736520e1 Peptidase Inhibitors
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marion Merrell Dow Inc.
 ; STREET: 2110 East Galbraith Rd.
 ; CITY: Cincinnati P. O. Box 156300
 ; STATE: Ohio
 ; COUNTRY: USA
 ; ZIP: 45215-6300
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/434,959
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/750,439
 ; FILING DATE: 20-AUG-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/454,803
 ; FILING DATE: 21-DEC-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/439,201
 ; FILING DATE: 20-NOV-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/416,817
 ; FILING DATE: 04-OCT-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/254,762
 ; FILING DATE: 07-OCT-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Nesbitt, Stephen L.
 ; REGISTRATION NUMBER: 28,981
 ; REFERENCE/DOCKET NUMBER: M01368F US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (513) 948-7965
 ; TELEFAX: (513) 948-7961
 ; TELEX: 214320
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-434-959-3

Query Match 100.0%; Score 21; DB 1; Length 4;

Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
 ||||
 Db 1 AAPF 4

RESULT 16
 US-08-434-959-4
 ; Sequence 4, Application US/08434959
 ; Patent No. 5736520
 ; GENERAL INFORMATION:
 ; APPLICANT: Bey, Philippe
 ; APPLICANT: Angelastro, Michael R
 ; APPLICANT: Mehdi, Shujaath
 ; TITLE OF INVENTION: No. 5736520e1 Peptidase Inhibitors
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marion Merrell Dow Inc.
 ; STREET: 2110 East Galbraith Rd.
 ; CITY: Cincinnati P. O. Box 156300
 ; STATE: Ohio
 ; COUNTRY: USA
 ; ZIP: 45215-6300
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/434,959
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/750,439
 ; FILING DATE: 20-AUG-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/454,803
 ; FILING DATE: 21-DEC-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/439,201
 ; FILING DATE: 20-NOV-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/416,817
 ; FILING DATE: 04-OCT-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/254,762
 ; FILING DATE: 07-OCT-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Nesbitt, Stephen L.
 ; REGISTRATION NUMBER: 28,981
 ; REFERENCE/DOCKET NUMBER: M01368F US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (513) 948-7965
 ; TELEFAX: (513) 948-7961
 ; TELEX: 214320
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-434-959-4

Query Match 100.0%; Score 21; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
 ||||
 Db 1 AAPF 4

RESULT 17
US-08-434-959-5
; Sequence 5, Application US/08434959
; Patent No. 5736520
; GENERAL INFORMATION:
; APPLICANT: Bey, Philippe
; APPLICANT: Angelastro, Michael R
; APPLICANT: Mehdi, Shujaath
; TITLE OF INVENTION: No. 5736520e1 Peptidase Inhibitors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marion Merrell Dow Inc.
; STREET: 2110 East Galbraith Rd.
; CITY: Cincinnati P. O. Box 156300
; STATE: Ohio
; COUNTRY: USA
; ZIP: 45215-6300
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,959
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/750,439
; FILING DATE: 20-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/454,803
; FILING DATE: 21-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/439,201
; FILING DATE: 20-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/416,817
; FILING DATE: 04-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/254,762
; FILING DATE: 07-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Nesbitt, Stephen L
; REGISTRATION NUMBER: 28,981
; REFERENCE/DOCKET NUMBER: M01368F US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 948-7965
; TELEFAX: (513) 948-7961
; TELEX: 214320
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-434-959-5

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 18
US-08-434-959-6
; Sequence 6, Application US/08434959
; Patent No. 5736520
; GENERAL INFORMATION:
; APPLICANT: Bey, Philippe

APPLICANT: Angelastro, Michael R
; APPLICANT: Mehdi, Shujaath
; TITLE OF INVENTION: No. 5736520e1 Peptidase Inhibitors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marion Merrell Dow Inc.
; STREET: 2110 East Galbraith Rd.
; CITY: Cincinnati P. O. Box 156300
; STATE: Ohio
; COUNTRY: USA
; ZIP: 45215-6300
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,959
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/750,439
; FILING DATE: 20-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/454,803
; FILING DATE: 21-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/439,201
; FILING DATE: 20-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/416,817
; FILING DATE: 04-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/254,762
; FILING DATE: 07-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Nesbitt, Stephen L
; REGISTRATION NUMBER: 28,981
; REFERENCE/DOCKET NUMBER: M01368F US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 948-7965
; TELEFAX: (513) 948-7961
; TELEX: 214320
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-434-959-6

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 19
US-08-460-343B-56
; Sequence 56, Application US/08460343B
; Patent No. 5741664
; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco

```
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,343B
FILING DATE: 01-Jun-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/398028
FILING DATE: 03-mar-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kubinec, Jeffrey S.
REGISTRATION NUMBER: 36,575
REFERENCE/DOCKET NUMBER: P0936C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8228
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-460-343B-56

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 4;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 20
US-08-067-180B-2
; Sequence 2, Application US/08067180B
; Patent No. 5750650
; GENERAL INFORMATION:
; APPLICANT: NAKANISHI, Koichiro
; APPLICANT: NOMURA, Keiichi
; APPLICANT: TAJIMA, Kyoko
; APPLICANT: HIRAYANI, Hajime
; APPLICANT: KATO, Kazuo
; TITLE OF INVENTION: FIBRINOLYTIC PROTEIN AND PRODUCTION
; TITLE OF INVENTION: METHOD THEREOF
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burgess, Ryan and Wayne
; STREET: 370 Lexington Avenue, Suite 2105
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5 1/4 INCH DISKETTE
; COMPUTER: PC'S LIMITED SYSTEM 200
; OPERATING SYSTEM: DOS
; SOFTWARE: WORDPERECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/067,180B
; FILING DATE: 24 - MAY 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/618,494
; FILING DATE: 27 - NOVEMBER 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: WAYNE, MILTON J.
; REGISTRATION NUMBER: 17,906
```

```
REFERENCE/DOCKET NUMBER: U-WP-4696CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-683-8150
TELEFAX: 212-532-4285
TELEX: 423794
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4
TYPE: AMINO ACID
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
US-08-067-180B-2

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 4;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 21
US-08-398-028B-56
; Sequence 56, Application US/08398028B
; Patent No. 5780285
; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/398,028B
; FILING DATE: 03-mar-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-398-028B-56

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 4;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 22
US-08-470-837-3
; Sequence 3, Application US/08470837
```

Patent No. 5800811
GENERAL INFORMATION:
APPLICANT: Nimml, Marcel E.
APPLICANT: Hall, Frederick L.
APPLICANT: Tuan, Tai-Lan
APPLICANT: Wu, Lingtao
APPLICANT: Cheung, David T.
TITLE OF INVENTION: Transforming Growth Factor B Fusion
TITLE OF INVENTION: and
TITLE OF INVENTION: Their Use In Wound Healing
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 11150 Santa Monica Boulevard, Suite 400
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90025-3395
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,837
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharp, Janice A.
REGISTRATION NUMBER: 34,051
REFERENCE/DOCKET NUMBER: 30630-1US01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310-445-1140
TELEFAX: 310-445-9031
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-470-837-3

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
Db 1 AAPF 4

RESULT 23
US-08-722-268-2
Sequence 2, Application US/08722268
Patent No. 5827662
GENERAL INFORMATION:
APPLICANT: Harvey Rubin, Barry Cooperman, No. 5827662man Schechter,
APPLICANT: Michael Plotkin, Zhi Wang
TITLE OF INVENTION: Methods of Producing Effective
TITLE OF INVENTION: Recombinant Serine Protease Inhibitors and Uses of These
TITLE OF INVENTION: Inhibitors
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jane Massey Licata, Esq.
STREET: 210 Lake Drive East, Suite 201
CITY: Cherry Hill
STATE: NJ
COUNTRY: USA
ZIP: 08002
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 486

OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,268
FILING DATE: December 18, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 276,936
FILING DATE: July 19, 1994
APPLICATION NUMBER: 221,078
FILING DATE: March 31, 1994
APPLICATION NUMBER: 221,171
FILING DATE: March 31, 1994
APPLICATION NUMBER: 005,908
FILING DATE: January 15, 1993
APPLICATION NUMBER: 735,335
FILING DATE: July 24, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: PENN-0107
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 779-2400
TELEFAX: (609) 779-8488
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4
TYPE: Amino Acid
TOPOLOGY: linear
US-08-722-268-2

Query Match 100.0%; Score 21; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
Db 1 AAPF 4

RESULT 24
US-08-504-265B-56
Sequence 56, Application US/08504265B
Patent No. 5837516
GENERAL INFORMATION:
APPLICANT: Marcus D. Ballinger and James A. Wells
TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
TITLE OF INVENTION: SUBSTRATES CONTAINING BASIC RESIDUES
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/504,265B
FILING DATE: 19-Jul-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/398028
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kubinec, Jeffrey S.
REGISTRATION NUMBER: 36,575
REFERENCE/DOCKET NUMBER: P0936P1
TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-8228
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-504-265B-56

Query Match 100.0%; Score 21; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
DB 1 AAPF 4

RESULT 25
US-08-439-005-10
Sequence 10, Application US/08439005
Patent No. 5843895

GENERAL INFORMATION:
APPLICANT: Lazarus, Robert A.
APPLICANT: Dennis, Mark S.
APPLICANT: Ulmer, Jana S.
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439, 005
FILING DATE: 11-MAY-1995

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/319501
FILING DATE: 04-OCT-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/121004
FILING DATE: 14-SEP-1993

ATTORNEY/AGENT INFORMATION:
NAME: Kubinec, Jeffrey S.
REGISTRATION NUMBER: 36,575
REFERENCE/DOCKET NUMBER: P0859C1D1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8228
TELEFAX: 415/952-9881
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-439-005-10

Query Match 100.0%; Score 21; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
DB 1 AAPF 4

US-08-439-005-10

RESULT 26
US-08-698-575E-2
Sequence 2, Application US/08698575E
Patent No. 5874585

GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: SUBSTITUTED HETEROCYCLIC COMPOUNDS USEFUL AS
INHIBITORS OF (SERINE PROTEASES) HUMAN NEUTROPHIL ELASTASE
NUMBER OF SEQUENCES: 4
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/698, 575E
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/345, 820
FILING DATE:

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-698-575E-2

Query Match 100.0%; Score 21; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
DB 1 AAPF 4

RESULT 27
US-08-424-017B-19
Sequence 19, Application US/08424017B
Patent No. 5935854

GENERAL INFORMATION:
APPLICANT: Foster, Donald C
APPLICANT: Sprecher, Cindy

TITLE OF INVENTION: NOVEL HUMAN AMYLOID PROTEIN PRECURSOR
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424, 017B
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/155, 331
FILING DATE:
APPLICATION NUMBER: US 07/985, 692
FILING DATE: 02-DEC-1992

US-08-424-017B-19

US-08-424-017B-19

US-08-424-017B-19

591087X


```
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-684
; REFERENCE/DOCKET NUMBER: 92-21C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..4
; OTHER INFORMATION: /label= ALA-1
; OTHER INFORMATION: /note= "Amino terminal alanine residue is capped
; OTHER INFORMATION: with a succinyl group"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..4
; OTHER INFORMATION: /label= Phe-4
; OTHER INFORMATION: /note= "Carboxyl-terminal phenylalanine residue is
; OTHER INFORMATION: capped with p-nitroanil..."
; US-08-424-017B-19

Query Match          100.0%; Score 21; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 28
US-08-751-070B-4
; Sequence 4, Application US/08751070B
; Patent No. 5976859
; GENERAL INFORMATION:
; APPLICANT: Damodaran, Srinivasan
; APPLICANT: Han, Xiao-Qing
; TITLE OF INVENTION: DETERGENT-STABLE MICROBIAL PROTEASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dewilt Ross & Stevens S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: U.S.A.
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,070B
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: 09820.028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-831-2100
; TELEFAX: 608-831-2106
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-751-070B-4

Query Match          100.0%; Score 21; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 29
US-09-069-823-2
; Sequence 2, Application US/09069823
; Patent No. 6037325
; GENERAL INFORMATION:
; APPLICANT: Gyorkos, Albert C.
; APPLICANT: Spruce, Lyle W.
; TITLE OF INVENTION: SUBSTITUTED HETEROCYCLIC COMPOUNDS
; TITLE OF INVENTION: USEFUL AS INHIBITORS OF (SERINE PROTEASES) HUMAN NEUTROPHIL
; FILE REFERENCE: 20774.240087
; CURRENT APPLICATION NUMBER: US/09/069,823
; CURRENT FILING DATE: 1998-04-30
; EARLIER APPLICATION NUMBER: 08/345,820
; EARLIER FILING DATE: 1994-11-21
; EARLIER APPLICATION NUMBER: 08/698,575
; EARLIER FILING DATE: 1996-08-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Tetrapeptide
; US-09-069-823-2

Query Match          100.0%; Score 21; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 1 AAPF 4

RESULT 30
US-09-076-460-1
; Sequence 1, Application US/09076460
; Patent No. 6080564
; GENERAL INFORMATION:
; APPLICANT: Laustsen, Mads
; APPLICANT: Nielsson, Stig
; TITLE OF INVENTION: Selective Inactivation of Aspergillus
; TITLE OF INVENTION: Proteases (As Amended)
; FILE REFERENCE: 4609.204-US
; CURRENT APPLICATION NUMBER: US/09/076,460
; CURRENT FILING DATE: 1998-05-12
; EARLIER APPLICATION NUMBER: 1392/95
; EARLIER FILING DATE: 1995-12-07
; EARLIER APPLICATION NUMBER: PCT/DK96/00489
; EARLIER FILING DATE: 1996-11-25
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: substrate
; US-09-076-460-1
Query Match      100.0%; Score 21; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 AAPF 4
      ||||
Db      1 AAPF 4

RESULT 31
US-08-950-618-3
; Sequence 3, Application US/08950618
; Patent No. 6087325
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: Ali, Shaikat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; TITLE OF INVENTION: Methods of Liposomal Drug Delivery Using
; FILE REFERENCE: TLC 215A
; CURRENT APPLICATION NUMBER: US/08/950,618
; CURRENT FILING DATE: 1997-10-15
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
US-08-950-618-3

Query Match      100.0%; Score 21; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 AAPF 4
      ||||
Db      1 AAPF 4

RESULT 32
US-08-950-618-18
; Sequence 18, Application US/08950618
; Patent No. 6087325
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: Ali, Shaikat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; TITLE OF INVENTION: Methods of Liposomal Drug Delivery Using
; FILE REFERENCE: TLC 215A
; CURRENT APPLICATION NUMBER: US/08/950,618
; CURRENT FILING DATE: 1997-10-15
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides

; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-terminal succinyl group
US-08-950-618-18
Query Match      100.0%; Score 21; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 AAPF 4
      ||||
Db      1 AAPF 4

RESULT 33
US-08-950-618-31
; Sequence 31, Application US/08950618
; Patent No. 6087325
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: Ali, Shaikat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; TITLE OF INVENTION: Methods of Liposomal Drug Delivery Using
; FILE REFERENCE: TLC 215A
; CURRENT APPLICATION NUMBER: US/08/950,618
; CURRENT FILING DATE: 1997-10-15
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
US-08-950-618-31

Query Match      100.0%; Score 21; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 AAPF 4
      ||||
Db      1 AAPF 4

RESULT 34
US-09-120-365-101
; Sequence 101, Application US/09120365
; Patent No. 6103514
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/120,365
; CURRENT FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 101
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Unknown
```

```
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: substrate
US-09-120-365-101

Query Match          100.0%; Score 21; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
       1111
Db      1 AAPF 4

RESULT 35
US-08-319-501-10
; Sequence 10, Application US/08319501
; Patent No. 6113896
; GENERAL INFORMATION:
; APPLICANT: Lazarus, Robert A.
; APPLICANT: Dennis, Mark S.
; APPLICANT: Ulmer, Jana S.
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING ECOTIN AND HOMOLOGS THE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/319,501
; FILING DATE: 4-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/121
; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 004
; FILING DATE: (null)
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Daryl B.
; REGISTRATION NUMBER: 32,637
; REFERENCE/DOCKET NUMBER: 859C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1249
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-319-501-10

Query Match          100.0%; Score 21; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
       1111
Db      1 AAPF 4

RESULT 36
US-09-168-010-3
; Sequence 3, Application US/09168010A
; Patent No. 6143716
```

```
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: Ali, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; APPLICANT: Cabral-Lilly, Donna
; TITLE OF INVENTION: Liposomal Peptide-Lipid Conjugates and
; TITLE OF INVENTION: Delivery Using Same
; FILE REFERENCE: TLC 215B
; CURRENT APPLICATION NUMBER: US/09/168,010A
; CURRENT FILING DATE: 1998-10-07
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
US-09-168-010-3

Query Match          100.0%; Score 21; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPF 4
       1111
Db      1 AAPF 4
```

```
RESULT 37
US-09-168-010-18
; Sequence 18, Application US/09168010A
; Patent No. 6143716
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: Ali, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erukulla, Ravi K.
; APPLICANT: Cabral-Lilly, Donna
; TITLE OF INVENTION: Liposomal Peptide-Lipid Conjugates and
; TITLE OF INVENTION: Delivery Using Same
; FILE REFERENCE: TLC 215B
; CURRENT APPLICATION NUMBER: US/09/168,010A
; CURRENT FILING DATE: 1998-10-07
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-terminal succinyl group
US-09-168-010-18

Query Match          100.0%; Score 21; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPF 4
       1111
Db      1 AAPF 4
```

RESULT 38
US-09-168-010-31
; Sequence 31, Application US/09168010A
; Patent No. 6143716
; GENERAL INFORMATION:
; APPLICANT: The Liposome Company Inc.
; APPLICANT: Pak, Charles
; APPLICANT: Meers, Paul
; APPLICANT: Ali, Shaukat
; APPLICANT: Janoff, Andrew S.
; APPLICANT: Franklin, J. Craig
; APPLICANT: Erkulalla, Ravi K.
; APPLICANT: Cabral-Lilly, Donna
; TITLE OF INVENTION: Liposomal Peptide-Lipid Conjugates and
; TITLE OF INVENTION: Delivery Using Same
; FILE REFERENCE: T1C 215B
; CURRENT APPLICATION NUMBER: US/09/168,010A
; EARLIER FILING DATE: 1998-10-07
; EARLIER APPLICATION NUMBER: US 60/027,544
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Peptides
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-terminal carboxy sugar group
US-09-168-010-31
Query Match 100.0%; Score 21; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPF 4
Db 1 AAPF 4
RESULT 39
US-08-905-359A-25
; Sequence 25, Application US/08905359A
; Patent No. 6153410
; GENERAL INFORMATION:
; APPLICANT: Frances H. Arnold
; APPLICANT: Zhixin Shao
; APPLICANT: Joseph A. Affholter
; APPLICANT: Huimin Zhao
; APPLICANT: Lori Giver
; TITLE OF INVENTION: Recombination of Polynucleotide
; TITLE OF INVENTION: Sequences Using Defined or Random Primer Sequences
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,359A
; FILING DATE: August 4, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/041,666

; FILING DATE: March 25, 1997
; APPLICATION NUMBER: 60/045,211
; FILING DATE: April 30, 1997
; APPLICATION NUMBER: 60/046,256
; FILING DATE: May 12, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 330187-84
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-905-359A-25
Query Match 100.0%; Score 21; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPF 4
Db 1 AAPF 4
RESULT 40
US-09-353-556-25
; Sequence 25, Application US/09353556
; Patent No. 6177263
; GENERAL INFORMATION:
; APPLICANT: Frances H. Arnold
; APPLICANT: Zhixin Shao
; APPLICANT: Joseph A. Affholter
; APPLICANT: Huimin Zhao
; APPLICANT: Lori Giver
; TITLE OF INVENTION: Recombination of Polynucleotide
; TITLE OF INVENTION: Sequences Using Defined or Random Primer Sequences
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/353,556
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,359
; FILING DATE: August 4, 1997
; APPLICATION NUMBER: 60/041,666
; FILING DATE: March 25, 1997
; APPLICATION NUMBER: 60/045,211
; FILING DATE: April 30, 1997
; APPLICATION NUMBER: 60/046,256
; FILING DATE: May 12, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 330187-84
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000

TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 25;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-353-556-25

Query Match 100.0%; Score 21; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPF 4
Db 1 AAPF 4

Search completed: December 6, 2002, 13:31:50
Job time : 10.3333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:30:14 ; Search time 4.88889 Seconds
(without alignments)
13.289 Million cell updates/sec

Title: US-10-033-526-1
Perfect score: 21
Sequence: 1 AAPF 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------------------|-------------------|
| 1 | 21 | 100.0 | 4 | 9 US-10-041-006A-10 | Sequence 10, Appl |
| 2 | 21 | 100.0 | 4 | 10 US-09-757-908A-15 | Sequence 15, Appl |
| 3 | 21 | 100.0 | 4 | 10 US-09-117-380B-2 | Sequence 2, Appl |
| 4 | 21 | 100.0 | 4 | 10 US-09-935-744-1 | Sequence 1, Appl |
| 5 | 21 | 100.0 | 4 | 12 US-10-090-624-33 | Sequence 33, Appl |
| 6 | 21 | 100.0 | 4 | 12 US-10-036-371-7 | Sequence 7, Appl |
| 7 | 21 | 100.0 | 4 | 12 US-10-040-655-10 | Sequence 10, Appl |
| 8 | 21 | 100.0 | 4 | 12 US-10-033-526-1 | Sequence 1, Appl |
| 9 | 21 | 100.0 | 6 | 10 US-09-994-927-1 | Sequence 1, Appl |
| 10 | 21 | 100.0 | 6 | 12 US-10-033-526-4 | Sequence 4, Appl |
| 11 | 21 | 100.0 | 10 | 9 US-10-068-965-4 | Sequence 4, Appl |
| 12 | 21 | 100.0 | 12 | 10 US-09-985-157-11 | Sequence 11, Appl |
| 13 | 21 | 100.0 | 25 | 10 US-09-864-761-38169 | Sequence 38169, A |
| 14 | 21 | 100.0 | 45 | 10 US-09-864-761-48747 | Sequence 48747, A |
| 15 | 21 | 100.0 | 51 | 10 US-09-925-297-566 | Sequence 566, App |
| 16 | 21 | 100.0 | 62 | 10 US-09-864-761-42500 | Sequence 42500, A |
| 17 | 21 | 100.0 | 75 | 10 US-09-864-761-38957 | Sequence 38957, A |
| 18 | 21 | 100.0 | 85 | 10 US-09-925-300-1841 | Sequence 1841, Ap |
| 19 | 21 | 100.0 | 89 | 10 US-09-925-301-1206 | Sequence 1206, Ap |

| | | | | | | |
|----|----|-------|-----|----|---------------------|-------------------|
| 20 | 21 | 100.0 | 100 | 10 | US-09-893-737-4 | Sequence 4, Appl |
| 21 | 21 | 100.0 | 113 | 9 | US-09-764-868-986 | Sequence 986, App |
| 22 | 21 | 100.0 | 125 | 10 | US-09-738-973-59 | Sequence 59, Appl |
| 23 | 21 | 100.0 | 132 | 10 | US-09-739-907-99 | Sequence 99, Appl |
| 24 | 21 | 100.0 | 135 | 10 | US-09-908-322-51 | Sequence 51, Appl |
| 25 | 21 | 100.0 | 137 | 10 | US-09-740-288A-2 | Sequence 2, Appl |
| 26 | 21 | 100.0 | 141 | 9 | US-09-736-457-1822 | Sequence 1822, Ap |
| 27 | 21 | 100.0 | 141 | 9 | US-09-902-941-1822 | Sequence 1822, Ap |
| 28 | 21 | 100.0 | 148 | 10 | US-09-916-790-30 | Sequence 30, Appl |
| 29 | 21 | 100.0 | 158 | 10 | US-09-925-300-1092 | Sequence 1092, Ap |
| 30 | 21 | 100.0 | 166 | 10 | US-09-925-301-1170 | Sequence 1170, Ap |
| 31 | 21 | 100.0 | 172 | 10 | US-09-739-907-87 | Sequence 87, Appl |
| 32 | 21 | 100.0 | 178 | 12 | US-10-052-586-512 | Sequence 512, App |
| 33 | 21 | 100.0 | 182 | 10 | US-09-739-907-191 | Sequence 191, App |
| 34 | 21 | 100.0 | 184 | 10 | US-09-764-864-1280 | Sequence 1280, Ap |
| 35 | 21 | 100.0 | 206 | 10 | US-09-815-242-10522 | Sequence 10522, A |
| 36 | 21 | 100.0 | 209 | 10 | US-09-811-284-225 | Sequence 225, App |
| 37 | 21 | 100.0 | 211 | 10 | US-09-925-302-652 | Sequence 652, App |
| 38 | 21 | 100.0 | 215 | 10 | US-09-820-893-67 | Sequence 67, Appl |
| 39 | 21 | 100.0 | 216 | 10 | US-09-815-242-5294 | Sequence 5294, Ap |
| 40 | 21 | 100.0 | 217 | 10 | US-09-815-242-10051 | Sequence 10051, A |
| 41 | 21 | 100.0 | 217 | 10 | US-09-815-242-14109 | Sequence 14109, A |
| 42 | 21 | 100.0 | 225 | 10 | US-09-905-810-1 | Sequence 1, Appl |
| 43 | 21 | 100.0 | 228 | 10 | US-09-815-242-10832 | Sequence 10832, A |
| 44 | 21 | 100.0 | 231 | 10 | US-09-815-242-12547 | Sequence 12547, A |
| 45 | 21 | 100.0 | 241 | 8 | US-08-450-842-5 | Sequence 5, Appl |

ALIGNMENTS

RESULT 1
US-10-041-006A-10
; Sequence 10, Application US/10041006A
; Patent No. US20020168754A1
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/10/041, 006A
; CURRENT FILING DATE: 2002-01-07
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide substrate
US-10-041-006A-10

Query Match 100.0%; Score 21; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
DB 1 AAPF 4

RESULT 2
US-09-757-908A-15
; Sequence 15, Application US/09757908A
; Patent No. US20020052468A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell
; TITLE OF INVENTION: Disulfide Core Polypeptides
; FILE REFERENCE: 98-13D1
; CURRENT APPLICATION NUMBER: US/09/757, 908A
; CURRENT FILING DATE: 2001-01-10

; PRIOR APPLICATION NUMBER: US 09/326,039
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: US 60/088,136
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-908A-15

Query Match 100.0%; Score 21; DB 10; Length 4;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AAPF 4
Db 1 AAPF 4

RESULT 3
US-09-117-380B-2
; Sequence 2, Application US/09117380B
; Patent No. US20020119917A1
; GENERAL INFORMATION:
; APPLICANT: FRIDKIN, Matityahu
; APPLICANT: YAVIN, Eran J.
; TITLE OF INVENTION: ANTI-INFLAMMATORY PEPTIDES DERIVED FROM C-REACTIVE
; FILE REFERENCE: FRIDKIN-1
; CURRENT APPLICATION NUMBER: US/09/117,380B
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: PCT/IL97/00032
; PRIOR FILING DATE: 1997-01-27
; PRIOR APPLICATION NUMBER: IL 116976
; PRIOR FILING DATE: 1996-01-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: The N-terminal Ala residue is modified with a
; OTHER INFORMATION: succinyl group; the C-terminal Phe residue is
; OTHER INFORMATION: modified with a nitroanilide group.
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-117-380B-2

Query Match 100.0%; Score 21; DB 10; Length 4;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AAPF 4
Db 1 AAPF 4

RESULT 4
US-09-935-744-1
; Sequence 1, Application US/09935744
; Patent No. US20020137118A1
; GENERAL INFORMATION:
; APPLICANT: Inouye, Masayori
; APPLICANT: Shinde, Ujwal
; APPLICANT: Fu, Xuan
; TITLE OF INVENTION: Biologically Active Protein Folding Intermediates
; FILE REFERENCE: 266/223
; CURRENT APPLICATION NUMBER: US/09/935,744
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: The sequence is part of a synthetic peptide that is used as a
; OTHER INFORMATION: strate for determining the activation time of a stable crossl
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(4)
; OTHER INFORMATION: Sequence is preceded by N-succinyl and followed by p-nitroanil
US-09-935-744-1

Query Match 100.0%; Score 21; DB 10; Length 4;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AAPF 4
Db 1 AAPF 4

RESULT 5
US-10-090-624-33
; Sequence 33, Application US/10090624
; Patent No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc_feature
; OTHER INFORMATION: Residue 1 is modified by a succinyl group.
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Residue 4 is modified by a p-nitroaniline group.
US-10-090-624-33

Query Match 100.0%; Score 21; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AAPF 4
Db 1 AAPF 4

RESULT 6
US-10-036-371-7
; Sequence 7, Application US/10036371
; Patent No. US20020141987A1
; GENERAL INFORMATION:
; APPLICANT: BJARNARSON, JON B.
; TITLE OF INVENTION: FISH SERINE PROTEINASES AND THEIR PHARMACEUTICAL AND
; TITLE OF INVENTION: COSMETIC USE

```
; FILE REFERENCE: 81691/284960
; CURRENT APPLICATION NUMBER: US/10/036,371
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/411,688
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 5086/99
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 7
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
US-10-036-371-7

Query Match      100.0%; Score 21; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
        ||||
Db      1 AAPF 4

RESULT 7
US-10-040-655-10
; Sequence 10, Application US/10040655
; Patent No. US20020146805A1
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/10/040,655
; CURRENT FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 10
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide substrate
US-10-040-655-10

Query Match      100.0%; Score 21; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
        ||||
Db      1 AAPF 4

RESULT 8
US-10-033-526-1
; Sequence 1, Application US/10033526
; Patent No. US20020147999A1
; GENERAL INFORMATION:
; APPLICANT: Robert W. Mahley
; APPLICANT: Yadong Huang
; TITLE OF INVENTION: Methods of Treating Disorders Related to
; FILE REFERENCE: UCAL217
; CURRENT APPLICATION NUMBER: US/10/033,526
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 60/245,737
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSEQ for windows version 4.0
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-033-526-1

Query Match      100.0%; Score 21; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
        ||||
Db      1 AAPF 4

RESULT 9
US-09-994-927-1
; Sequence 1, Application US/09994927
; Patent No. US20020127605A1
; GENERAL INFORMATION:
; APPLICANT: Guilford Pharmaceuticals Inc.
; APPLICANT: Hamilton, Gregory
; APPLICANT: Belyakov, Sergei
; APPLICANT: Vaal, Mark
; APPLICANT: Wei, Ling
; APPLICANT: Wu, Yong-Qian
; APPLICANT: Steiner, Joseph
; TITLE OF INVENTION: Bisubstituted Carbocyclic Cyclophilin Binding Compounds and Th
; FILE REFERENCE: 03166.0029.NPUS02
; CURRENT APPLICATION NUMBER: US/09/994,927
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,074
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/291,966
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 1
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Model substrate for measuring rotamase inhibition activity
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa at position 1 is N-succinyl
; NAME/KEY: MISC_FEATURE
; LOCATION: (6)..(6)
; OTHER INFORMATION: Xaa at position 6 is p-nitroanillide
US-09-994-927-1

Query Match      100.0%; Score 21; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
        ||||
Db      2 AAPF 5

RESULT 10
US-10-033-526-4
; Sequence 4, Application US/10033526
; Patent No. US20020147999A1
; GENERAL INFORMATION:
; APPLICANT: Robert W. Mahley
; APPLICANT: Yadong Huang
; TITLE OF INVENTION: Methods of Treating Disorders Related to
; FILE REFERENCE: UCAL217
```

```
; CURRENT APPLICATION NUMBER: US/10/033,526
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 60/245,737
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
; US-10-033-526-4
```

```
Query Match      100.0%; Score 21; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. NO. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AAPF 4
        ||||
Db       3 AAPF 6
```

```
RESULT 11
US-10-068-965-4
; Sequence 4, Application US/10068965
; Patent No. US20020156123A1
; GENERAL INFORMATION:
; APPLICANT: BALLIGNAD, JEAN-LUC
; APPLICANT: FERON, OLIVIER
; TITLE OF INVENTION: NOVEL PHARMACEUTICAL COMPOSITIONS FOR MODULATING
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: DCLERC-2 P1
; CURRENT APPLICATION NUMBER: US/10/068,965
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: PCT/EP00/07731
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: 99870171
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Caveolin
; US-10-068-965-4
```

```
Query Match      100.0%; Score 21; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. NO. 11;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AAPF 4
        ||||
Db       3 AAPF 6
```

```
RESULT 12
US-09-985-157-11
; Sequence 11, Application US/09985157
; Patent No. US20020146797A1
; GENERAL INFORMATION:
; APPLICANT: SHOKAT, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases Which Can Utilize Modified Nucleotide
; TITLE OF INVENTION: Triphosphate Substrates
; FILE REFERENCE: 51538-5002-05
; CURRENT APPLICATION NUMBER: US/09/985,157
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US98/02522
```

```
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: US 08/797,552
; PRIOR FILING DATE: 1997-02-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Optimized Ab1 substrate
; US-09-985-157-11
```

```
Query Match      100.0%; Score 21; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. NO. 14;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AAPF 4
        ||||
Db       5 AAPF 8
```

```
RESULT 13
US-09-864-761-38169
; Sequence 38169, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
```

```
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38169
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL050342.36
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.95
US-09-864-761-38169
```

```
Query Match          100.0%; Score 21; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPF 4
        ||||
Db       2 AAPF 5
```

```
RESULT 14
US-09-864-761-48747
; Sequence 48747, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
```

```
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48747
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006389.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
; OTHER INFORMATION: SWISSPROT HIT: P15772, EVALU8.20e+00
US-09-864-761-48747
```

```
Query Match          100.0%; Score 21; DB 10; Length 45;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPF 4
        ||||
Db       39 AAPF 42
```

```
RESULT 15
US-09-925-297-566
; Sequence 566, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 566
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (35)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (36)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (48)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-566

Query Match          100.0%; Score 21; DB 10; Length 51;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPF 4
        ||||
Db       21 AAPF 24
```

```
RESULT 16
US-09-864-761-42500
; Sequence 42500, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
```



```
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42500
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005323.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.5
; OTHER INFORMATION: EST_HUMAN HIT: AA777043.1, EVALUE 6.00e-21
; OTHER INFORMATION: SWISSPROT HIT: P13535, EVALUE 2.00e-25
US-09-864-761-42500

Query Match      100.0%; Score 21; DB 10; Length 62;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAPF 4
        ||||
Db      13 AAPF 16
```

```
RESULT 17
US-09-864-761-38957
; Sequence 38957, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
```

```
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecmlca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38957
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004636.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.4
; OTHER INFORMATION: SWISSPROT HIT: P07774, EVALUE 5.20e+00
; OTHER INFORMATION: EST_HUMAN HIT: AA469992.1, EVALUE 4.00e-03
US-09-864-761-38957
```

```
Query Match      100.0%; Score 21; DB 10; Length 75;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAPF 4
        ||||
Db      12 AAPF 15
```

```
RESULT 18
US-09-925-300-1841
; Sequence 1841, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1841
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1841

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 22 AAPF 25

RESULT 19
US-09-925-301-1206
; Sequence 1206, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1206
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1206

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 78 AAPF 81

RESULT 20
US-09-893-737-4
; Sequence 4, Application US/09893737

; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-737-4

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 100;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 71 AAPF 74

RESULT 21
US-09-764-868-986
; Sequence 986, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 986
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (62)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (77)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (80)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (90)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-986

Query Match
Best Local Similarity 100.0%; Score 21; DB 9; Length 113;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 8 AAPF 11

RESULT 22
US-09-738-973-59
```

```
; Sequence 59, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-59
```

```
Query Match          100.0%; Score 21; DB 10; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAPF 4
        ||||
Db       9 AAPF 12
```

```
RESULT 23
US-09-739-907-99
; Sequence 99, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 99
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (132)
; OTHER INFORMATION: Xaa equals stop translation
US-09-739-907-99
```

```
Query Match          100.0%; Score 21; DB 10; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAPF 4
        ||||
Db       28 AAPF 31
```

```
RESULT 24
US-09-908-322-51
; Sequence 51, Application US/09908322
; Patent No. US20020107194A1
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos Manuel Pinto
; APPLICANT: Lewis, Julian Hart
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Gray, Grace
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF
; TITLE OF INVENTION: VERTEBRATE DELTA GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/908,322
; FILING DATE: 17-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/981,392
; FILING DATE: 22-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 135 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-09-908-322-51
```

```
Query Match          100.0%; Score 21; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAPF 4
        ||||
Db       34 AAPF 37
```

```
RESULT 25
US-09-740-288A-2
; Sequence 2, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
```

```
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/09/740,288A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Hordeum vulgare
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (131)..(131).
; OTHER INFORMATION: xaa = any amino acid
US-09-740-288A-2

Query Match          100.0%; Score 21; DB 10; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
    ||||
Db 23 AAPF 26

RESULT 26
US-09-736-457-1822
; Sequence 1822, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1822
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-1822

Query Match          100.0%; Score 21; DB 9; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
    ||||
Db 25 AAPF 28

RESULT 27
US-09-902-941-1822
; Sequence 1822, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.
; APPLICANT: Marnierakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1822
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-941-1822

Query Match          100.0%; Score 21; DB 9; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
    ||||
Db 25 AAPF 28

RESULT 28
US-09-916-790-30
; Sequence 30, Application US/09916790
; Patent No. US20020061573A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachael
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 18431 AND 32374, NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 381552002700
; CURRENT APPLICATION NUMBER: US/09/916,790
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,543
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-916-790-30

Query Match          100.0%; Score 21; DB 10; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
    ||||
Db 48 AAPF 51

RESULT 29
US-09-925-300-1092
; Sequence 1092, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
```

```
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1092
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1092

Query Match          100.0%; Score 21; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
       1111
Db      50 AAPF 53

RESULT 30
US-09-925-301-1170
; Sequence 1170, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1170
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (18)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (131)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1170

Query Match          100.0%; Score 21; DB 10; Length 166;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
       1111
Db      31 AAPF 34

RESULT 31
US-09-739-907-87
; Sequence 87, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
```

```
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 87
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (170)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-739-907-87

Query Match          100.0%; Score 21; DB 10; Length 172;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
       1111
Db      28 AAPF 31

RESULT 32
US-10-052-586-512
; Sequence 512, Application US/10052586
; Patent No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052,586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
```


Fri Dec 6 14:08:47 2002

us-10-033-526-1.mod.rapb

Page 11

| | |
|----|-------------------------------------|
| 1 | PRIOR FILING DATE: 1997-10-29 |
| 2 | PRIOR APPLICATION NUMBER: 60/063870 |
| 3 | PRIOR FILING DATE: 1997-10-31 |
| 4 | PRIOR APPLICATION NUMBER: 60/064103 |
| 5 | PRIOR FILING DATE: 1997-10-31 |
| 6 | PRIOR APPLICATION NUMBER: 60/065311 |
| 7 | PRIOR FILING DATE: 1997-11-13 |
| 8 | PRIOR APPLICATION NUMBER: 60/066120 |
| 9 | PRIOR FILING DATE: 1997-11-21 |
| 10 | PRIOR APPLICATION NUMBER: 60/066466 |
| 11 | PRIOR FILING DATE: 1997-11-24 |
| 12 | PRIOR APPLICATION NUMBER: 60/066772 |
| 13 | PRIOR FILING DATE: 1997-11-24 |
| 14 | PRIOR APPLICATION NUMBER: 60/069335 |
| 15 | PRIOR FILING DATE: 1997-12-11 |
| 16 | PRIOR APPLICATION NUMBER: 60/069425 |
| 17 | PRIOR FILING DATE: 1997-12-12 |
| 18 | PRIOR APPLICATION NUMBER: 60/069870 |
| 19 | PRIOR FILING DATE: 1997-12-17 |
| 20 | PRIOR APPLICATION NUMBER: 60/068017 |
| 21 | PRIOR FILING DATE: 1997-12-18 |
| 22 | PRIOR APPLICATION NUMBER: 60/077450 |
| 23 | PRIOR FILING DATE: 1998-03-10 |
| 24 | PRIOR APPLICATION NUMBER: 60/077632 |
| 25 | PRIOR FILING DATE: 1998-03-11 |
| 26 | PRIOR APPLICATION NUMBER: 60/077649 |
| 27 | PRIOR FILING DATE: 1998-03-11 |
| 28 | PRIOR APPLICATION NUMBER: 60/078886 |
| 29 | PRIOR FILING DATE: 1998-03-20 |
| 30 | PRIOR APPLICATION NUMBER: 60/078939 |
| 31 | PRIOR FILING DATE: 1998-03-20 |
| 32 | PRIOR APPLICATION NUMBER: 60/079664 |
| 33 | PRIOR FILING DATE: 1998-03-27 |
| 34 | PRIOR APPLICATION NUMBER: 60/079786 |
| 35 | PRIOR FILING DATE: 1998-03-27 |
| 36 | PRIOR APPLICATION NUMBER: 60/080107 |
| 37 | PRIOR FILING DATE: 1998-03-31 |
| 38 | PRIOR APPLICATION NUMBER: 60/080194 |
| 39 | PRIOR FILING DATE: 1998-03-31 |
| 40 | PRIOR APPLICATION NUMBER: 60/080327 |
| 41 | PRIOR FILING DATE: 1998-04-01 |
| 42 | PRIOR APPLICATION NUMBER: 60/080333 |
| 43 | PRIOR FILING DATE: 1998-04-01 |
| 44 | PRIOR APPLICATION NUMBER: 60/081049 |
| 45 | PRIOR FILING DATE: 1998-04-08 |
| 46 | PRIOR APPLICATION NUMBER: 60/081070 |
| 47 | PRIOR FILING DATE: 1998-04-08 |
| 48 | PRIOR APPLICATION NUMBER: 60/081195 |
| 49 | PRIOR FILING DATE: 1998-04-09 |
| 50 | PRIOR APPLICATION NUMBER: 60/081838 |
| 51 | PRIOR FILING DATE: 1998-04-15 |
| 52 | PRIOR APPLICATION NUMBER: 60/082568 |
| 53 | PRIOR FILING DATE: 1998-04-21 |
| 54 | PRIOR APPLICATION NUMBER: 60/082569 |
| 55 | PRIOR FILING DATE: 1998-04-21 |
| 56 | PRIOR APPLICATION NUMBER: 60/082704 |
| 57 | PRIOR FILING DATE: 1998-04-22 |
| 58 | PRIOR APPLICATION NUMBER: 60/082797 |
| 59 | PRIOR FILING DATE: 1998-04-22 |
| 60 | PRIOR APPLICATION NUMBER: 60/083322 |
| 61 | PRIOR FILING DATE: 1998-04-28 |
| 62 | PRIOR APPLICATION NUMBER: 60/083495 |
| 63 | PRIOR FILING DATE: 1998-04-29 |
| 64 | PRIOR APPLICATION NUMBER: 60/083496 |
| 65 | PRIOR FILING DATE: 1998-04-29 |
| 66 | PRIOR APPLICATION NUMBER: 60/083499 |
| 67 | PRIOR FILING DATE: 1998-04-29 |
| 68 | PRIOR APPLICATION NUMBER: 60/083559 |
| 69 | PRIOR FILING DATE: 1998-04-29 |
| 70 | PRIOR APPLICATION NUMBER: 60/084366 |
| 71 | PRIOR FILING DATE: 1998-05-05 |
| 72 | PRIOR APPLICATION NUMBER: 60/084414 |
| 73 | PRIOR FILING DATE: 1998-05-06 |

| | |
|----|--------------------------------------|
| 1 | PRIOR APPLICATION NUMBER: 60/084639 |
| 2 | PRIOR FILING DATE: 1998-05-07 |
| 3 | PRIOR APPLICATION NUMBER: 1998-05-07 |
| 4 | PRIOR FILING DATE: 1998-05-07 |
| 5 | PRIOR APPLICATION NUMBER: 60/084640 |
| 6 | PRIOR FILING DATE: 1998-05-07 |
| 7 | PRIOR APPLICATION NUMBER: 60/084643 |
| 8 | PRIOR FILING DATE: 1998-05-07 |
| 9 | PRIOR APPLICATION NUMBER: 60/085573 |
| 10 | PRIOR FILING DATE: 1998-05-15 |
| 11 | PRIOR APPLICATION NUMBER: 60/085579 |
| 12 | PRIOR FILING DATE: 1998-05-15 |
| 13 | PRIOR APPLICATION NUMBER: 60/085580 |
| 14 | PRIOR FILING DATE: 1998-05-15 |
| 15 | PRIOR APPLICATION NUMBER: 60/085582 |
| 16 | PRIOR FILING DATE: 1998-05-15 |
| 17 | PRIOR APPLICATION NUMBER: 60/085700 |
| 18 | PRIOR FILING DATE: 1998-05-15 |
| 19 | PRIOR APPLICATION NUMBER: 60/086023 |
| 20 | PRIOR FILING DATE: 1998-05-18 |
| 21 | PRIOR APPLICATION NUMBER: 60/086392 |
| 22 | PRIOR FILING DATE: 1998-05-22 |
| 23 | PRIOR APPLICATION NUMBER: 60/086486 |
| 24 | PRIOR FILING DATE: 1998-05-22 |
| 25 | PRIOR APPLICATION NUMBER: 60/087098 |
| 26 | PRIOR FILING DATE: 1998-05-28 |
| 27 | PRIOR APPLICATION NUMBER: 60/087208 |
| 28 | PRIOR FILING DATE: 1998-05-28 |
| 29 | PRIOR APPLICATION NUMBER: 60/087609 |
| 30 | PRIOR FILING DATE: 1998-06-02 |
| 31 | PRIOR APPLICATION NUMBER: 60/087759 |
| 32 | PRIOR FILING DATE: 1998-06-02 |
| 33 | PRIOR APPLICATION NUMBER: 60/087827 |
| 34 | PRIOR FILING DATE: 1998-06-03 |
| 35 | PRIOR APPLICATION NUMBER: 60/088025 |
| 36 | PRIOR FILING DATE: 1998-06-04 |
| 37 | PRIOR APPLICATION NUMBER: 60/088028 |
| 38 | PRIOR FILING DATE: 1998-06-04 |
| 39 | PRIOR APPLICATION NUMBER: 60/088029 |
| 40 | PRIOR FILING DATE: 1998-06-04 |
| 41 | PRIOR APPLICATION NUMBER: 60/088033 |
| 42 | PRIOR FILING DATE: 1998-06-04 |
| 43 | PRIOR APPLICATION NUMBER: 60/088167 |
| 44 | PRIOR FILING DATE: 1998-06-05 |
| 45 | PRIOR APPLICATION NUMBER: 60/088202 |
| 46 | PRIOR FILING DATE: 1998-06-05 |
| 47 | PRIOR APPLICATION NUMBER: 60/088212 |
| 48 | PRIOR FILING DATE: 1998-06-05 |
| 49 | PRIOR APPLICATION NUMBER: 60/088217 |
| 50 | PRIOR FILING DATE: 1998-06-05 |
| 51 | PRIOR APPLICATION NUMBER: 60/088326 |
| 52 | PRIOR FILING DATE: 1998-06-04 |
| 53 | PRIOR APPLICATION NUMBER: 60/088655 |
| 54 | PRIOR FILING DATE: 1998-06-09 |
| 55 | PRIOR APPLICATION NUMBER: 60/088722 |
| 56 | PRIOR FILING DATE: 1998-06-10 |
| 57 | PRIOR APPLICATION NUMBER: 60/088738 |
| 58 | PRIOR FILING DATE: 1998-06-10 |
| 59 | PRIOR APPLICATION NUMBER: 60/088740 |
| 60 | PRIOR FILING DATE: 1998-06-10 |
| 61 | PRIOR APPLICATION NUMBER: 60/088811 |
| 62 | PRIOR FILING DATE: 1998-06-10 |
| 63 | PRIOR APPLICATION NUMBER: 60/088824 |
| 64 | PRIOR FILING DATE: 1998-06-10 |
| 65 | PRIOR APPLICATION NUMBER: 60/088825 |
| 66 | PRIOR FILING DATE: 1998-06-10 |
| 67 | PRIOR APPLICATION NUMBER: 60/088826 |
| 68 | PRIOR FILING DATE: 1998-06-10 |
| 69 | PRIOR APPLICATION NUMBER: 60/088861 |
| 70 | PRIOR FILING DATE: 1998-06-11 |
| 71 | PRIOR APPLICATION NUMBER: 60/088863 |
| 72 | PRIOR FILING DATE: 1998-06-11 |
| 73 | PRIOR APPLICATION NUMBER: 60/088876 |
| 74 | PRIOR FILING DATE: 1998-06-11 |
| 75 | PRIOR APPLICATION NUMBER: 60/089090 |

```
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089908
```

```
Query Match          100.0%; Score 21; DB 12; Length 178;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPF 4
        ||||
Db       85 AAPF 88
```

RESULT 33

```
US-09-739-907-191
; Sequence 191, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 191
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (180)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-739-907-191
```

```
Query Match          100.0%; Score 21; DB 10; Length 182;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPF 4
        ||||
Db       38 AAPF 41
```

RESULT 34

```
US-09-764-864-1280
; Sequence 1280, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
```

```
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1280
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (35)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1280
```

```
Query Match          100.0%; Score 21; DB 10; Length 184;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPF 4
        ||||
Db       21 AAPF 24
```

RESULT 35

```
US-09-815-242-10522
; Sequence 10522, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10522
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10522
```

```
Query Match          100.0%; Score 21; DB 10; Length 206;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPF 4
        ||||
Db       42 AAPF 45
```

```
RESULT 36
US-09-811-284-225
; Sequence 225, Application US/09811284
; Patent No. US20020058306A1
; GENERAL INFORMATION:
; APPLICANT: Vogel, Gabriel
; TITLE OF INVENTION: No. US20020058306A1 G Protein-Coupled Receptors
; FILE REFERENCE: 00167US1
; CURRENT APPLICATION NUMBER: US/09/811,284
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,783
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,907
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,918
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,960
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,917
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/192,945
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,916
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,923
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,933
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,830
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,234
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,155
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,935
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 225
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-284-225

Query Match      100.0%; Score 21; DB 10; Length 209;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
        ||||
Db      16 AAPF 19

RESULT 37
US-09-925-302-652
; Sequence 652, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 652
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-09-925-302-652

Query Match      100.0%; Score 21; DB 10; Length 211;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
        ||||
Db      194 AAPF 197

RESULT 38
US-09-820-893-67
; Sequence 67, Application US/09820893
; Patent No. US20020076705A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 31 Human Secreted Proteins
; FILE REFERENCE: PZ033P1
; CURRENT APPLICATION NUMBER: US/09/820,893
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/531,119
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/102,895
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 67
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-820-893-67

Query Match      100.0%; Score 21; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
        ||||
Db      45 AAPF 48

RESULT 39
US-09-815-242-5294
; Sequence 5294, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
```

```

; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5294
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5294

```

```

Query Match          100.0%; Score 21; DB 10; Length 216;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 AAPF 4
      ||||
Db      99 AAPF 102

```

```

RESULT 40
US-09-815-242-10051
; Sequence 10051, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10051
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10051

```

```

Query Match          100.0%; Score 21; DB 10; Length 217;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 AAPF 4
      ||||
Db      100 AAPF 103

```

Search completed: December 6, 2002, 13:42:05
 Job time : 5.88889 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:25:59 ; Search time 10.2222 Seconds
(Without alignments)
37.618 Million cell updates/sec

Title: US-10-033-526-1
Perfect score: 21
Sequence: 1 AAPF 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|---------------------|
| 1 | 21 | 100.0 | 5 | 2 | JS0319 | subesophageal gang |
| 2 | 21 | 100.0 | 42 | 2 | S43963 | hypothetical prote |
| 3 | 21 | 100.0 | 47 | 2 | PN0607 | cytochrome-c oxida |
| 4 | 21 | 100.0 | 50 | 2 | AD3575 | hypothetical prote |
| 5 | 21 | 100.0 | 63 | 1 | OSB08A | cytochrome-c oxida |
| 6 | 21 | 100.0 | 63 | 2 | SI0303 | cytochrome-c oxida |
| 7 | 21 | 100.0 | 75 | 2 | AB2778 | hypothetical prote |
| 8 | 21 | 100.0 | 79 | 2 | C69125 | hypothetical prote |
| 9 | 21 | 100.0 | 80 | 2 | G96953 | probable metal-bin |
| 10 | 21 | 100.0 | 86 | 2 | E90905 | hypothetical prote |
| 11 | 21 | 100.0 | 88 | 2 | F97429 | flilQ protein (AF30 |
| 12 | 21 | 100.0 | 88 | 2 | AG2647 | flagellar biosynth |
| 13 | 21 | 100.0 | 90 | 2 | S02854 | male accessory gla |
| 14 | 21 | 100.0 | 90 | 2 | S30428 | Mst26ab protein - |
| 15 | 21 | 100.0 | 90 | 2 | S30412 | male accessory gla |
| 16 | 21 | 100.0 | 90 | 2 | S30430 | Mst26ab protein - |
| 17 | 21 | 100.0 | 94 | 2 | C98239 | hypothetical prote |
| 18 | 21 | 100.0 | 97 | 2 | AB1502 | probable cell surf |
| 19 | 21 | 100.0 | 100 | 2 | G86174 | hypothetical prote |
| 20 | 21 | 100.0 | 101 | 2 | D90761 | hypothetical prote |
| 21 | 21 | 100.0 | 102 | 2 | B86976 | PE-family protein |
| 22 | 21 | 100.0 | 102 | 2 | H70898 | probable PE protei |
| 23 | 21 | 100.0 | 107 | 2 | H75397 | hypothetical prote |
| 24 | 21 | 100.0 | 108 | 2 | F72507 | hypothetical prote |
| 25 | 21 | 100.0 | 109 | 2 | G86433 | protein T17H7.7 [I |
| 26 | 21 | 100.0 | 112 | 2 | H96031 | probable transcrip |
| 27 | 21 | 100.0 | 122 | 2 | PC4279 | anti-SS-A/Ro 60K p |
| 28 | 21 | 100.0 | 122 | 2 | PC4280 | anti-SS-A/Ro 60K p |
| 29 | 21 | 100.0 | 122 | 2 | C83501 | hypothetical prote |

| | | | | | | |
|----|----|-------|-----|---|--------|--------------------|
| 30 | 21 | 100.0 | 125 | 2 | AD0605 | probable membrane |
| 31 | 21 | 100.0 | 125 | 2 | A83048 | hypothetical prote |
| 32 | 21 | 100.0 | 126 | 2 | T05908 | probable ribosomal |
| 33 | 21 | 100.0 | 127 | 2 | A25864 | calcitonin gene-re |
| 34 | 21 | 100.0 | 127 | 2 | A71334 | hypothetical prote |
| 35 | 21 | 100.0 | 128 | 1 | TCHUR | calcitonin gene-re |
| 36 | 21 | 100.0 | 129 | 2 | I37779 | ig variable region |
| 37 | 21 | 100.0 | 130 | 2 | T29498 | hypothetical prote |
| 38 | 21 | 100.0 | 132 | 2 | S36196 | hypothetical prote |
| 39 | 21 | 100.0 | 133 | 2 | S76285 | hypothetical prote |
| 40 | 21 | 100.0 | 133 | 2 | F75297 | hypothetical prote |
| 41 | 21 | 100.0 | 133 | 2 | S75430 | hypothetical prote |
| 42 | 21 | 100.0 | 134 | 2 | A44173 | calcitonin gene-re |
| 43 | 21 | 100.0 | 134 | 2 | A72489 | hypothetical prote |
| 44 | 21 | 100.0 | 137 | 2 | F87627 | conserved hypothet |
| 45 | 21 | 100.0 | 139 | 2 | D75542 | hypothetical prote |

ALIGNMENTS

RESULT 1
JS0319
subesophageal ganglion pentapeptide - house cricket
C:Species: Acheta domestica (house cricket)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000
C:Accession: JS0319
R:Wicker, C.; Wicker, C.
Comp. Biochem. Physiol. C 88, 185-187, 1987
A:Title: Isolation and structure of a peptide isolated from the subesophageal gangli
A:Reference number: JS0319
A:Accession: JS0319
A:Molecule type: protein
A:Residues: 1-5 <WIC>

Query Match 100.0%; Score 21; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 2 AAPF 5

RESULT 2
S43963
hypothetical protein (clone pR57) - Rhizobium sp. (strain NGR234) (fragment)
C:Species: Rhizobium sp.
A:Variety: strain NGR234
C:Date: 20-Oct-1994 #sequence_revision 23-Feb-1996 #text_change 21-Aug-1998
C:Accession: S43963
R:Perret, X.; Fellay, R.; Bjourson, A.J.; Cooper, J.E.; Brenner, S.; Broughton, W.J.
Nucleic Acids Res. 22, 1335-1341, 1994
A:Title: Subtraction hybridisation and shot-gun sequencing: a new approach to identif
A:Reference number: S43961; MUID:94248027; PMID:8190622
A:Accession: S43963
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-42 <PER>
A:Experimental source: strain NGR234
C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match 100.0%; Score 21; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 10 AAPF 13

RESULT 3
PN0607


```

cytochrome-c oxidase (EC 1.9.3.1) chain VIIC - pig (fragment)
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C/Accession: PN0607
R/Sillard, R.; Joenvall, H.; Mutl, V.
Biochem. Biophys. Res. Commun. 195, 746-750, 1993
A/Title: Characterization of porcine intestinal cytochrome c oxidase subunit VIIC, purified from pig
A/Reference number: PN0607; MUID:93384597; PMID:8396926
A/Accession: PN0607
A/Molecule type: protein
A/Residues: 1-47 <SIL>
A/Experimental source: intestine
C/Superfamily: cytochrome-c oxidase chain VIIC
C/Keywords: mitochondrion; oxidoreductase; respiratory chain

Query Match          100.0%; Score 21; DB 2; length 47;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1 AAPF 4
        ||||
Db       34 AAPF 37

```

```

RESULT 4
AD3575
hypothetical protein BMEI10525 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AD3575
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AD3575
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-50 <KUR>
A:Cross-references: GB:AE008918; PIDN:AAL53767.1; PID:gl7984695; GSPDB:GN00191
C:Genetics:
A:Gene: BMEI10525
A:Map position: II

Query Match          100.0%;   Score 21;   DB 2;   Length 50;
Best Local Similarity 100.0%;   Pred. NO. 1.2e+02;
Matches 4;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

OY      1 AAFP 4
        |||
Db      44 AAFP 47

RESULT 5
OSB08A
cytochrome-c oxidase (EC 1.9.3.1) chain VIIC precursor [validated] - bovine
N:Alternate names: cytochrome-c oxidase chain VIIIA
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-May-1979 #sequence_revision 31-Dec-1992 #text_change 15-Sep-2000
C:Accession: JH0473; S06597; A00498; H29968; S18834
R:Aqua, M.S.; Bachman, N.J.; Lomax, M.I.; Grossman, L.I.
Gene 104, 211-217, 1991
A:Title: Characterization and expression of a cDNA specifying subunit VIIC of bovine cy
A:Reference number: JH0473; MUID:92009215; PMID:1655579
A:Accession: JH0473
A:Molecule type: DNA
A:Residues: 1-63 <AOU1>
A:Cross-references: GB:X58823
R:Aqua, M.S.; Lomax, M.I.; Schon, E.A.; Grossman, L.I.
Nucleic Acids Res. 17, 8376, 1989
A:Title: Nucleotide sequence of a cDNA for bovine cytochrome c oxidase subunit VIIC.
A:Reference number: S06597; MUID:90045968; PMID:2554257

```

A:Accession: S06597
A:Molecule type: mRNA
A:Residues: 1-63 <AOU2>
A:Cross-references: EMBL:X15725; NID:g277; PIDN:CAB57793.1; PID:g6015488
R:Buse, G.; Steffens, G.J.
Hoppe-Seyler's Z. Physiol. Chem. 359, 1005-1009, 1978
A:Title: Studies on cytochrome c oxidase, II. The chemical constitution of a short polypeptide chain from bovine heart mitochondria
A:Reference number: A00498; MUID:79046803; PMID:213363
A:Accession: A00498
A:Molecule type: protein
A:Residues: 17-63 <BUS>
A:Experimental source: heart
R:Yanamura, W.; Zhang, Y.Z.; Takamiya, S.; Capaldi, R.A.
Biochemistry 27, 4909-4914, 1988
A:Title: Tissue-specific differences between heart and liver cytochrome c oxidase.
A:Reference number: A90531; MUID:89000697; PMID:2844245
A:Accession: H29968
A:Molecule type: protein
A:Residues: 17-42 <YAN>
A:Experimental source: liver
R:Tsuchihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Ito
submitted to the Brookhaven Protein Data Bank, April 1996
A:Reference number: A67451; PDB:1OCC
A:Contents: annotation; X-ray crystallography, 2.8 angstroms, residues 17-63
R:Tsuchihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Ito
Science 272, 1136-1144, 1996
A:Title: The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 Å resolution
A:Reference number: A57981; MUID:96216288; PMID:8638158
A:Contents: annotation; X-ray crystallography, 2.8 angstroms
C:Genetics:
A:Genome: nuclear
C:Complex: part of a 13 chain complex spanning the inner mitochondrial membrane and cytochrome c (see PIR:CABO), Vb (see PIR:OGBO6A), Via (see PIR:OGBO6), VId (see PIR:OGBO7), ViC (see PIR:OGBO8)mers within the mitochondrial inner-membrane
C:Function:
A:Description: the cytochrome-c oxidase complex catalyzes the oxidation of four molecules from the mitochondrial matrix producing two molecules of water and lowering the potential across the membrane
A:Pathway: oxidative phosphorylation; respiratory chain
A>Note: the role of chain VIIC is not clear
C:Superfamily: cytochrome-c oxidase chain VIIC
C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane
F:1-16/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F:17-63/Product: cytochrome-c oxidase chain VIIC #status experimental <MAT>
F:17-33/Domain: mitochondrial matrix #status experimental <MML>
F:34-60/Domain: transmembrane helix #status experimental <TR01>
F:61-63/Domain: intracrystal #status experimental <ITC1>

| Query Match | Best Local Similarity | Score 21; | DB 1; | Length 63; |
|-------------|-----------------------|---------------|-----------|------------|
| Matches 4; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

| | | | |
|----|----|------|----|
| QY | 1 | AAPF | 4 |
| | | | |
| Db | 50 | AAPF | 53 |

RESULT 6

S10303

cytochrome-c oxidase (EC 1.9.3.1) chain VIIC - mouse

C:Species: Mus musculus (house mouse)

C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 11-Jun-1999

C:Accession: S10303

R:Akamatsu, M.; Grossman, L.I.
Nucleic Acids Res. 18, 3645, 1990

A:Title: Nucleotide sequence of a cDNA for mouse cytochrome c oxidase subunit VIIC.

A:Reference number: S10303; MUID:90301494; PMID:2163523

A:Accession: S10303

A:Molecule type: mRNA

A:Residues: 1-63 <AKA>

A:Cross-references: EMBL:X52940; NID:g50524; PIDN:CAA37115.1; PID:g50525

C:Genetics:

A:Genome: nuclear

C:Superfamily: cytochrome-c oxidase chain VIIC

C:Keywords: mitochondrion; oxidoreductase

Query Match 100.0%; Score 21; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
1111
Db 50 AAPF 53

RESULT 7

AB2778
hypothetical protein Atu1639 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AB2778
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; Mclellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AB2778
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-75 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL42640.1; PID:g17740071; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu1639
A:Map position: circular chromosome

Query Match 100.0%; Score 21; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
1111
Db 50 AAPF 53

RESULT 8

C69125
hypothetical protein MTH206 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: C69125
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: C69125
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-79 <MTH>
A:Cross-references: GB:AE000807; GB:AE000666; NID:g2621239; PIDN:AB84712.1; PID:g262125
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH206
A:Start codon: TTG

Query Match 100.0%; Score 21; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
1111
Db 36 AAPF 39

RESULT 9
G96953
probable metal-binding protein [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: G96953
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L.
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: G96953
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-80 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK78418.1; PID:g15023293; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0438

Query Match 100.0%; Score 21; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
1111
Db 63 AAPF 66

RESULT 10

E90905
hypothetical protein ECS2213 [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: E90905
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E90905
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA835636.1; PID:g13361679; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECS2213

Query Match 100.0%; Score 21; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
1111
Db 26 AAPF 29

RESULT 11

F97429
f11Q protein (AF300968) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: F97429
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: F97429
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-88 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK86391.1; PID:g15155521; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_1023
A:Map position: circular chromosome

Query Match 100.0%; Score 21; DB 2; Length 88;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
||||
Db 65 AAPF 68

RESULT 12

AG2647
flagellar biosynthetic protein [imported] - Agrobacterium tumefaciens (strain C58, Dupon
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AG2647
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AG2647
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-88 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL41597.1; PID:g17738933; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: fljQ
A:Map position: circular chromosome

Query Match 100.0%; Score 21; DB 2; Length 88;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
||||
Db 65 AAPF 68

RESULT 13

S02854
male accessory gland secretory protein 26Ab precursor - fruit fly (Drosophila melanogast
N:Alternate names: male accessory gland secretory protein MSP355b; male paragonial prote
C:Species: Drosophila melanogaster
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 21-Jul-2000
C:Accession: S02854; S30408; S30416; S30420; S30426; S30410; S30424
R:Monsma, S.A.; Wolfiner, M.F.
Genes Dev. 2, 1063-1073, 1988
A:Title: Structure and expression of a Drosophila male accessory gland gene whose produc
A:Reference number: S02853; MUID:89053045; PMID:3142802
A:Accession: S02854
A:Molecule type: DNA
A:Residues: 1-90 <MON>
A:Cross-references: EMBL:Y00219; NID:g8264; PIDN:CAA68367.1; PID:g8266
A:Experimental source: strain Canton-S
R:Aguade, M.; Miyashita, N.; Langley, C.H.
Genetics 132, 755-770, 1992
A:Title: Polymorphism and divergence in the Mst26A male accessory gland gene region in D
A:Reference number: S30407; MUID:93106377; PMID:1361475
A:Accession: S30408
A:Molecule type: DNA
A:Residues: 1-90 <AGU>
A:Cross-references: EMBL:X70888; NID:g3402845; PIDN:CAA50233.1; PID:g8234

A:Experimental source: allele NC1
A:Accession: S30416
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-90 <AGW>

A:Cross-references: EMBL:X70892; NID:g8244; PID:g8246

A:Experimental source: allele NC5

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

A:Accession: S30420

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-90 <AGF>

A:Cross-references: EMBL:X70894; NID:g8250; PID:g8252

A:Experimental source: allele NC7

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

A:Accession: S30426

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-90 <AGA>

A:Cross-references: EMBL:X70897; NID:g8259; PID:g8260

A:Experimental source: allele NC10

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

A:Accession: S30410

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-79, 'V', 81-90 <AG3>

A:Cross-references: EMBL:X70889; NID:g3402846; PIDN:CAA50235.1; PID:g8237

A:Experimental source: allele NC2

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

A:Accession: S30424

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-63, 'H', 65-90 <AG4>

A:Cross-references: EMBL:X70896; NID:g3406838; PIDN:CAA50249.1; PID:g8258

A:Experimental source: allele NC9

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

C:Genetics:

A:Gene: FlyBase:Acp26Ab

A:Cross-references: FlyBase:FBgn0002856

A:Map position: 2

A:Introns: 11/1

C:Superfamily: male accessory gland secretory protein 26Ab

F:1-18/Domain: signal sequence #status predicted <Sig>

F:19-90/Product: male accessory gland secretory protein 26Ab #status predicted <Mat>

Query Match 100.0%; Score 21; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
||||
Db 21 AAPF 24

RESULT 14

S30428
Mst26Ab protein - fruit fly (Drosophila mauritiana)
C:Species: Drosophila mauritiana
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C:Accession: S30428
R:Aguade, M.; Miyashita, N.; Langley, C.H.
Genetics 132, 755-770, 1992
A:Title: Polymorphism and divergence in the Mst26A male accessory gland gene region 1
A:Reference number: S30407; MUID:93106377; PMID:1361475
A:Accession: S30428
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-90 <AGU>
A:Cross-references: EMBL:X70898; NID:g8261; PIDN:CAA50253.1; PID:g8263
C:Genetics:
A:Gene: FlyBase:Dmau/Acp26Ab
A:Cross-references: FlyBase:FBgn0012495
A:Introns: 11/1

C:Superfamily: male accessory gland secretory protein 26Ab

Query Match 100.0%; Score 21; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
| | | |
Db 21 AAPF 24

RESULT 15

S30412

male accessory gland secretory protein 26Ab precursor (variant 1) - fruit fly (*Drosophila*)
N:Alternate names: male accessory gland secretory protein 355b; Mst26Ab protein
C:Species: *Drosophila melanogaster*

C:Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
C:Accession: S30412; S30414; S30418; S30422
R:Aguade, M.; Miyashita, N.; Langley, C.H.
Genetics 132, 755-770, 1992

A:Title: Polymorphism and divergence in the Mst26A male accessory gland gene region in *D.*
A:Reference number: S30407; MUID:93106377; PMID:1361475

A:Accession: S30412

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-90 <AGU>

A:Cross-references: EMBL:X70890; NID:g3402847; PIDN:CAA50237.1; PID:g8240

A:Experimental source: allele NC3

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

A:Accession: S30414

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-90 <AGW>

A:Cross-references: EMBL:X70891; NID:g8241; PID:g8243

A:Experimental source: allele NC4

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

A:Accession: S30418

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-90 <AGF>

A:Cross-references: EMBL:X70893; NID:g8247; PID:g8249

A:Experimental source: allele NC6

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

A:Accession: S30422

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-90 <AGA>

A:Cross-references: EMBL:X70895; NID:g8253; PID:g8255

A:Experimental source: allele NC8

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

C:Genetics:

A:Gene: FlyBase:Acp26Aa

A:Cross-references: FlyBase:FBgn0002855

A:Map position: 2

A:Introns: 11/1

C:Superfamily: male accessory gland secretory protein 26Ab

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-90/Product: male accessory gland secretory protein 26Ab #status predicted <MAT>

Query Match 100.0%; Score 21; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
| | | |
Db 21 AAPF 24

RESULT 16

S30430

Mst26Ab protein - fruit fly (*Drosophila simulans*)
C:Species: *Drosophila simulans*

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-Jan-2000

C:Accession: S30430
R:Aguade, M.; Miyashita, N.; Langley, C.H.
Genetics 132, 755-770, 1992

A:Title: Polymorphism and divergence in the Mst26A male accessory gland gene region i
A:Reference number: S30407; MUID:93106377; PMID:1361475

A:Accession: S30430

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-90 <AGU>

A:Cross-references: EMBL:X70899; NID:g9134; PID:g9136

C:Genetics:

A:Gene: FlyBase:Dsim/Acp26Ab

A:Cross-references: FlyBase:FBgn0012821

A:Introns: 11/1

C:Superfamily: male accessory gland secretory protein 26Ab

Query Match 100.0%; Score 21; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
| | | |
Db 21 AAPF 24

RESULT 17

C98239

hypothetical protein AGR_L1727 [imported] - *Agrobacterium tumefaciens* (strain C58, C
C:Species: *Agrobacterium tumefaciens*

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002

C:Accession: C98239

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium*

A:Reference number: A97359; PMID:11743194

A:Accession: C98239

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-94 <KUR>

A:Cross-references: GB:AE007870; PIDN:AAK89437.1; PID:g15159298; GSPDB:GN00170

C:Genetics:

A:Gene: AGR_L1727

A:Map position: linear chromosome

Query Match 100.0%; Score 21; DB 2; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
| | | |
Db 16 AAPF 19

RESULT 18

AB1502

probable cell surface protein (LPXTG motif) [imported] - *Listeria innocua* (strain C11

C:Species: *Listeria innocua*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AB1502

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
. ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tjelle, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1502

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-97 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC95786.1; PID:g16412994; GSPDB:GN00178

A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lln0554

Query Match 100.0%; Score 21; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
||||
Db 16 AAPF 19

RESULT 19

G86174

hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: G86174

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G86174

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-100 <STO>

A:Cross-references: GB:AE005172; NID:g2341040; PIDN:AAB70444.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 100.0%; Score 21; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
||||
Db 81 AAPF 84

RESULT 20

D90761

hypothetical protein ECs1060 [imported] - Escherichia coli (strain O157:H7, substrain R1

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001

C:Accession: D90761

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: D90761

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-101 <HAY>

A:Cross-references: GB:BA000007; PIDN:BAB34483.1; PID:g13360520; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain R1MD 0509952

C:Genetics:

A:Gene: ECs1060

Query Match 100.0%; Score 21; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
||||

Db 26 AAPF 29

RESULT 21

B86976

PE-family protein [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: B86976

R:Coie, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; R.; Davies, R.M.; Devlin, K.; Dutfoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holro eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;

A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: B86976

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-102 <STO>

A:Cross-references: GB:AL450380; NID:g13092742; PIDN:CAC30046.1; GSPDB:GN00147

C:Genetics:

A:Gene: PE

Query Match 100.0%; Score 21; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
||||
Db 31 AAPF 34

RESULT 22

H70898

probable PE protein - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: H70898

R:Coie, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: H70898

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-102 <COL>

A:Cross-references: GB:Z80108; GB:AL123456; NID:g3256012; PIDN:CAB02191.1; PID:e26555

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: PE

Query Match 100.0%; Score 21; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
||||
Db 31 AAPF 34

RESULT 23

H75397

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: H75397

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J ; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: H75397
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-107 <WHI>
A:Cross-references: GB:AE001986; GB:AE000513; NID:96459162; PIDN:AAF10985.1; PID:9645917
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1405
A:Map position: 1

Query Match 100.0%; Score 21; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
|||
Db 78 AAPF 81

RESULT 24
F72507
hypothetical protein APE2036 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C:Accession: F72507
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain K1
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: F72507
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <KAW>
A:Cross-references: DDBJ:AP000063; NID:95105654; PIDN:BAA81046.1; PID:91044832; PID:9510
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2036
C:Superfamily: Aeropyrum pernix hypothetical protein APE2036

Query Match 100.0%; Score 21; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
|||
Db 95 AAPF 98

RESULT 25
G86433
protein T17H7.7 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G86433
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G86433
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-109 <STO>

A:Cross-references: GB:AE005172; NID:94926822; PIDN:AAD32932.1; GSPDB:GN00141
C:Genetics:
A:Gene: T17H7.7
A:Map position: 1

Query Match 100.0%; Score 21; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
|||
Db 70 AAPF 73

RESULT 26
H96031
probable transcription regulator, ArsR family protein [imported] - *Sinorhizobium meliloti*
C:Species: *Sinorhizobium meliloti*
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: H96031
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: H96031
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-112 <RUB>
A:Cross-references: GB:AL591985; PIDN:CAC49920.1; PID:915141408; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R. L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMD20608
A:Genome: plasmid

Query Match 100.0%; Score 21; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
|||
Db 34 AAPF 37

RESULT 27 -
PC4279
anti-SS-A/Ro 60K peptide heavy chain E-42 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 21-Jan-2000
C:Accession: PC4279
R:Suzuki, H.; Takemura, H.; Suzuki, M.; Sekine, Y.; Kashiwagi, H.
Biochem. Biophys. Res. Commun. 232, 101-106, 1997
A:Title: Molecular cloning of anti-SS-A/Ro 60-kDa peptide fab fragments from infiltra
A:Reference number: PC4279; MUID:97236289; PMID:9125110
A:Accession: PC4279
A:Molecule type: protein
A:Residues: 1-122 <SUZ>
C:Comment: This antibody is commonly found in systemic autoimmune diseases such as Sj
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 21; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
||||
Db 106 AAPF 109

RESULT 28

PC4280
anti-SS-A/Ro 60K peptide heavy chain E-60 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 21-Jan-2000
C:Accession: PC4280
R:Suzuki, H.; Takemura, H.; Suzuki, M.; Sekine, Y.; Kashiwagi, H.
Biochem. Biophys. Res. Commun. 232, 101-106, 1997
A:Title: Molecular cloning of anti-SS-A/Ro 60-kDa peptide fab fragments from infiltratin
A:Reference number: PC4279; MUID:97236289; PMID:9125110
A:Accession: PC4280
A:Molecule type: protein
A:Residues: 1-122 <SUZ>
C:Comment: This antibody is commonly found in systemic autoimmune diseases such as Sjog
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 21; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
||||
Db 106 AAPF 109

RESULT 29

C83501
hypothetical protein PA1149 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83501
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
. ; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: C83501
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-122 <STO>
A:Cross-references: GB:AE004545; GB:AE004091; NID:g99947070; PIDN:AAG04538.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1149

Query Match 100.0%; Score 21; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
||||
Db 14 AAPF 17

RESULT 30

AD0605
probable membrane protein SRY0904 [imported] - Salmonella enterica subsp. enterica serov
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AD0605
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conington, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
. S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

A:Reference number: AB0502; PMID:11677608

A:Accession: AD0605

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-125 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD05310.1; PID:gl6502074; GSPDB:GN00176

C:Genetics:

A:Gene: SRY0904

C:Superfamily: Escherichia coli probable membrane protein ybjM

Query Match 100.0%; Score 21; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
||||
Db 71 AAPF 74

RESULT 31

A83048
hypothetical protein PA4795 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: A83048
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
. ; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: A83048
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <STO>
A:Cross-references: GB:AE004892; GB:AE004091; NID:g99951049; PIDN:AAG08181.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4795

Query Match 100.0%; Score 21; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
||||
Db 25 AAPF 28

RESULT 32

T05908
probable ribosomal protein S8 - barley (fragment)
C:Species: Hordeum vulgare (barley)
C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 20-Jun-2000
C:Accession: T05908
R:Hess, W.R.; Golz, R.R.; Boerner, T.
Plant Sci. 133, 191-201, 1998
A:Title: Analysis of randomly selected cDNAs reveals the expression of stress- and d
A:Reference number: 215411
A:Accession: T05908
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-126 <HES>
A:Cross-references: EMBL:AJ000228; PIDN:CAA03954.1
A:Experimental source: cv. Haisa
C:Superfamily: rat ribosomal protein S8
C:Keywords: protein biosynthesis; ribosome

Query Match 100.0%; Score 21; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
1111
Db 107 AAPF 110

RESULT 33
A25864

calcitonin gene-related peptide beta precursor - human
N:Alternate names: calcitonin gene-related peptide II
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 29-Aug-1997
C:Accession: A25864; JH0620; B26142; A34565
R:Steenbergh, P.H.; Hoepfener, J.W.M.; Zandberg, J.; Visser, A.; Lips, C.J.M.; Jansz, H.
FEBS Lett. 209, 97-103, 1986
A:Title: Structure and expression of the human calcitonin/CGRP genes.
A:Reference number: A25864; MUID:87105923; PMID:3492393
A:Accession: A25864
A:Molecule type: DNA
A:Residues: 1-127 <STE>
R:Kitamura, K.; Kangawa, K.; Kawamoto, M.; Ichiki, Y.; Matsuo, H.; Eto, T.
Biochem. Biophys. Res. Commun. 185, 134-141, 1992
A:Title: Isolation and characterization of peptides which act on rat platelets, from a p
A:Reference number: JH0618; MUID:92287083; PMID:1318039
A:Accession: JH0620
A:Molecule type: protein
A:Residues: 82,'X','84-87','X','89-104 <KIT>
A:Experimental source: pheochromocytoma
R:Petermann, J.B.; Born, W.; Chang, J.Y.; Fischer, J.A.
J. Biol. Chem. 262, 542-545, 1987
A:Title: Identification in the human central nervous system, pituitary, and thyroid of a
A:Reference number: A92637; MUID:87109142; PMID:3492492
A:Accession: B26142
A:Molecule type: protein
A:Residues: 82,'X','84-87','X','89-91','X','93-98','X','100-105','X','107-109 <PET>
R:Mimalawansa, S.J.; Morris, H.R.; Etienne, A.; Blench, I.; Panico, M.; MacIntyre, I.
Biochem. Biophys. Res. Commun. 167, 993-1000, 1990
A:Title: Isolation, purification and characterization of beta-hCGRP from human spinal co
A:Reference number: A34565; MUID:90211348; PMID:2322288
A:Accession: A34565
A:Molecule type: protein
A:Residues: 82-86;104-117 <NIM>
C:Comment: Calcitonin gene-related peptide II peptide is a potent vasorelaxant.
C:Genetics:
A:Gene: GDB:CALCB; CALC2
A:Cross-references: GDB:120572; OMIM:114160
A:Map position: 11p15.2-11p15.1
C:Superfamily: calcitonin
C:Keywords: amidated carboxyl end; neuropeptide
F:82-118/Product: calcitonin gene-related peptide beta #status experimental <MAT>
F:83-88/Disulfide bonds: #status experimental
F:118/Modified site: amidated carboxyl end (Phe) (amide in mature form from following gl

Query Match 100.0%; Score 21; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
1111
Db 25 AAPF 28

RESULT 34
A71334
hypothetical protein TP0355 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C:Accession: A71334
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uterback, T.; McDo
they, L.; Weldman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876

A:Accession: A71334
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-127 <COL>
A:Cross-references: GB:AE001215; GB:AE000520; NID:93322631; PIDN:AAC65353.1; PID:9332
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0355
C:Superfamily: syphilis spirochete hypothetical protein TP0355

Query Match 100.0%; Score 21; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
1111
Db 85 AAPF 88

RESULT 35
TCHUR

calcitonin gene-related peptide alpha precursor [validated] - human
N:Alternate names: calcitonin gene-related peptide I; CGRP-I
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1987 #sequence_revision 02-Jul-1996 #text_change 08-Dec-2000
C:Accession: S07644; A22949; B22716; I55536; JF0005; S10813; A26142; JH0619; I52204;
R:Broad, P.M.; Symes, A.J.; Thakker, R.V.; Craig, R.K.
Nucleic Acids Res. 17, 6999-7011, 1989
A:Title: Structure and methylation of the human calcitonin/alpha-CGRP gene.
A:Reference number: S07643; MUID:89386053; PMID:25711128
A:Accession: S07644
A:Molecule type: DNA
A:Residues: 1-128 <BRO>
A:Cross-references: EMBL:X15943; NID:g29613; PIDN:CAA34070.1; PID:g296638
A:Note: the authors translated the codon CAG for residue 19 as Glu
R:Jonas, V.; Lin, C.R.; Kawashima, E.; Semion, D.; Swanson, L.W.; Mermod, J.J.; Evans,
Proc. Natl. Acad. Sci. U.S.A. 82, 1994-1998, 1985
A:Title: Alternative RNA processing events in human calcitonin/calcitonin gene-relate
A:Reference number: A94030; MUID:85166259; PMID:3872459
A:Accession: A22949
A:Molecule type: mRNA
A:Residues: 1-128 <JON>
A:Cross-references: GB:M12667; NID:g179825; PIDN:AAA51914.1; PID:g179828
R:Edbrooke, M.R.; Parker, D.; McVey, J.H.; Riley, J.H.; Sorenson, G.D.; Pettengill, O
EMBO J. 4, 715-724, 1985
A:Title: Expression of the human calcitonin/CGRP gene in lung and thyroid carcinoma.
A:Reference number: A91034; MUID:85230541; PMID:2408883
A:Accession: B22716
A:Molecule type: mRNA
A:Residues: 'V','50-75','S','76-128 <EDB>
R:Steenbergh, P.H.; Hopfener, J.W.; Zandberg, J.; Van de Ven, W.J.; Jansz, H.S.; Lips
J. Clin. Endocrinol. Metab. 59, 358-360, 1984
A:Title: Calcitonin gene related peptide coding sequence is conserved in the human ge
A:Reference number: I55536; MUID:84240176; PMID:6610687
A:Accession: I55536
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 77-128 <RES>
A:Cross-references: GB:M28637; NID:g180467; PIDN:AAA52012.1; PID:g457134
R:Morris, H.R.; Panico, M.; Etienne, T.; Tipples, J.; Girgis, S.I.; MacIntyre, I.
Nature 308, 746-748, 1984
A:Title: Isolation and characterization of human calcitonin gene-related peptide.
A:Reference number: A93329; MUID:84191466; PMID:6609312
A:Accession: JF0005
A:Molecule type: protein
A:Residues: 83-119 <MOR>
A:Note: this peptide was detected in medullary thyroid carcinoma tissue and in plasma
R:Zaidi, M.; Brain, S.D.; Tipples, J.R.; di Marzo, V.; Moonga, B.S.; Chambers, T.J.;
Biochem. J. 269, 775-780, 1990
A:Title: Structure-activity relationship of human calcitonin-gene-related peptide.
A:Reference number: S10813; MUID:90358780; PMID:2390067
A:Accession: S10813
A:Molecule type: protein

A:Residues: 83-99,'A',101-119 <ZAI>
R:Petermann, J.B.; Born, W.; Chang, J.Y.; Fischer, J.A.
J. Biol. Chem. 262, 542-545, 1987
A:Title: Identification in the human central nervous system, pituitary, and thyroid of a
A:Reference number: A92637; MUID:87109142; PMID:3492492
A:Accession: A26142
A:Molecule type: protein
A:Residues: 83-88,'X',90-101,'X',103-111,'X',113-115,'X',117 <PET>
R:Kitamura, K.; Kangawa, K.; Kawamoto, M.; Ichiki, Y.; Matsuo, H.; Eto, T.
Biochem. Biophys. Res. Commun. 185, 134-141, 1992
A:Title: Isolation and characterization of peptides which act on rat platelets, from a P
A:Reference number: JH0618; MUID:92287083; PMID:1318039
A:Accession: JH0619
A:Molecule type: protein
A:Residues: 83,'X',85-88,'X',90-108 <KIT>
A:Experimental source: pheochromocytoma
R:Neikin, B.D.; Rosenfeld, K.I.; de Bustros, A.; Leong, S.S.; Roos, B.A.; Baylin, S.B.
Biochem. Biophys. Res. Commun. 123, 648-655, 1984
A:Title: Structure and expression of a gene encoding human calcitonin and calcitonin gen
A:Reference number: I52204; MUID:85022523; PMID:6148938
A:Accession: I52204
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 48-119 <RE2>
A:Cross-references: GB:K03512; NID:g180465; PIDN:AAA52011.1; PID:g180466
R:Craig, R.K.; Riley, J.H.; Edbrooke, M.R.; Broad, P.M.; Foord, S.M.; Al-Kazwini, S.J.;
Biochem. Soc. Symp. 52, 91-105, 1986
A:Title: Expression and function of the human calcitonin/alpha-CGRP gene in health and d
A:Reference number: I39387; MUID:87213363; PMID:3034287
A:Accession: I84508
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 77-128 <RE3>
A:Cross-references: GB:M26094; NID:g179798; PIDN:AAA51912.1; PID:g179799
C:Comment: This peptide is a potent vasorelaxant.
C:Comment: This peptide increases the rate and force of contraction of rat auricles in v
C:Genetics:
A:Gene: GDB:CALCA; CALCI
A:Cross-references: GDB:120571; OMIM:114130
A:Map position: 11p15.2-11p15.1
A:Introns: 29/2; 76/2
C:Superfamily: calcitonin
C:Keywords: alternative splicing; amidated carboxyl end; neuropeptide; vasodilator
F:83-119/Product: calcitonin gene-related peptide alpha #status experimental <CTN>
F:84-89/Disulfide bonds: #status experimental
F:119/Modified site: amidated carboxyl end (Phe) (amide in mature form from following gl

Query Match 100.0%; Score 21; DB 1; Length 128;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
Db 25 AAPF 28

RESULT 36
I37779
Ig variable region (VDJ) (clone T20-24) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 21-Jan-2000
C:Accession: I37779; S25473
R:Demelson, C.; Chastagner, P.; Theze, J.; Zouali, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 514-518, 1994
A:Title: Somatic diversification in the heavy chain variable region genes expressed by h
A:Reference number: A36876; MUID:94119917; PMID:8290556
A:Accession: I37779
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-129 <RES>
A:Cross-references: EMBL:X67910; NID:g33576; PIDN:CAA48108.1; PID:g33577
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:23-105/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 21; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
Db 113 AAPF 116

RESULT 37
T29498
hypothetical protein K06B9.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29498
R:Miller, N.; Bradshaw, H.
submitted to the EMBL Data Library, February 1996
A:Description: The sequence of C. elegans cosmid K06B9.
A:Reference number: 220628
A:Accession: T29498
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-130 <MIL>
A:Cross-references: EMBL:U50072; PIDN:AAA93448.1; CESP:K06B9.3
C:Genetics:
A:Gene: CESP:K06B9.3
A:Introns: 32/3; 64/1

Query Match 100.0%; Score 21; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
Db 4 AAPF 7

RESULT 38
S36196
hypothetical protein 122.1 - Haloferax mediterranei
C:Species: Haloferax mediterranei
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: S36196; S42855; S69772
R:Mojica, F.J.M.; Juez, G.; Rodriguez-Valera, F.
Mol. Microbiol. 9, 613-621, 1993
A:Title: Transcription at different salinities of Haloferax mediterranei sequences ad
A:Reference number: S36196; MUID:94018655; PMID:8412707
A:Accession: S36196
A:Molecule type: DNA
A:Residues: 1-132 <MOJ>
A:Cross-references: EMBL:X73453; NID:g437778; PIDN:CAA51831.1; PID:g437779
R:Mojica, F.J.M.; Ferrer, C.; Juez, G.; Rodriguez-Valera, F.
Mol. Microbiol. 17, 85-93, 1995
A:Title: Long stretches of short tandem repeats are present in the largest replicons
A:Reference number: S69772; MUID:96020664; PMID:7476211
A:Contents: annotation
C:Superfamily: Haloferax mediterranei hypothetical protein 122.1

Query Match 100.0%; Score 21; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
Db 125 AAPF 128

RESULT 39
S76285
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C;Accession: S76285
 R;KaneKO, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 s.
 A;Reference number: S74322; MUID:97061201; PMID:8905231
 A;Accession: S76285
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-133 <KAN>
 A;Cross-references: EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BAAL0137.1; PID:g100151
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C;Superfamily: Escherichia coli protein P15B

Query Match 100.0%; Score 21; DB 2; Length 133;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
 ||||
 Db 73 AAPF 76

RESULT 40
 F75297
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C;Species: Deinococcus radiodurans
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C;Accession: F75297
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 , M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A;Reference number: A75250; MUID:20036896; PMID:10567266
 A;Accession: F75297
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-133 <WHI>
 A;Cross-references: GB:AE002057; GB:AE000513; NID:g6460050; PIDN:AAF11800.1; PID:g646005
 C;Genetics:
 A;Experimental source: strain R1
 A;Gene: DR2252
 A;Map position: 1

Query Match 100.0%; Score 21; DB 2; Length 133;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
 ||||
 Db 7 AAPF 10

Search completed: December 6, 2002, 13:31:01
 Job time : 10.2222 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:20:29 ; Search time 5.55556 Seconds
(without alignments)
29.863 Million cell updates/sec

Title: US-10-033-526-1
Perfect score: 21
Sequence: 1 AAPF 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|------------|--------------------|
| 1 | 21 | 100.0 | 5 | 1 | SUGA_ACHDO | P19991 acheta dome |
| 2 | 21 | 100.0 | 63 | 1 | COXO_BOVIN | P00430 bos taurus |
| 3 | 21 | 100.0 | 63 | 1 | COXO_MOUSE | P17665 mus musculu |
| 4 | 21 | 100.0 | 90 | 1 | MS2B_DROME | P33738 drosophila |
| 5 | 21 | 100.0 | 90 | 1 | MS2B_DROME | P10334 drosophila |
| 6 | 21 | 100.0 | 90 | 1 | MS2B_DROSE | P33739 drosophila |
| 7 | 21 | 100.0 | 90 | 1 | MS2B_DROSI | P33740 drosophila |
| 8 | 21 | 100.0 | 102 | 1 | U185_ORYSA | O94dm8 oryza sativ |
| 9 | 21 | 100.0 | 102 | 1 | YD86_MYCTU | P71656 mycobacteri |
| 10 | 21 | 100.0 | 127 | 1 | CAL2_HUMAN | P10092 homo sapien |
| 11 | 21 | 100.0 | 127 | 1 | Y355_TREPA | O83374 treponema p |
| 12 | 21 | 100.0 | 128 | 1 | CAL1_HUMAN | P06881 homo sapien |
| 13 | 21 | 100.0 | 130 | 1 | CAL2_MOUSE | O99mp3 mus musculu |
| 14 | 21 | 100.0 | 130 | 1 | SZ05_RAT | P97885 ratus norv |
| 15 | 21 | 100.0 | 133 | 1 | P2Y2_BOVIN | O18951 bos taurus |
| 16 | 21 | 100.0 | 133 | 1 | RBFA_SYNY3 | O55625 synechocyst |
| 17 | 21 | 100.0 | 134 | 1 | CAL2_RAT | P10093 ratus norv |
| 18 | 21 | 100.0 | 134 | 1 | CU3A_TENMO | P80683 tenebrio mo |
| 19 | 21 | 100.0 | 136 | 1 | NU2M_ARTSA | P19042 artemia sai |
| 20 | 21 | 100.0 | 141 | 1 | CAL0_HUMAN | P01258 homo sapien |
| 21 | 21 | 100.0 | 141 | 1 | YEF5_YEAST | P32616 saccharomyc |
| 22 | 21 | 100.0 | 142 | 1 | PSAH_MAIZE | O65101 zea mays (m |
| 23 | 21 | 100.0 | 143 | 1 | PSAH_HORVU | P20143 hordeum vul |
| 24 | 21 | 100.0 | 144 | 1 | PSAH_SPIOL | P22179 spinacia ol |
| 25 | 21 | 100.0 | 145 | 1 | PSAH_BRARA | O04006 brassica ra |
| 26 | 21 | 100.0 | 145 | 1 | PSH1_ARATH | O9su17 arabidopsis |
| 27 | 21 | 100.0 | 145 | 1 | PSH2_ARATH | O9su16 arabidopsis |
| 28 | 21 | 100.0 | 165 | 1 | VNS3_CVPPU | P09047 porcine tra |
| 29 | 21 | 100.0 | 168 | 1 | RBS_SACHY | O41373 saccharum h |
| 30 | 21 | 100.0 | 181 | 1 | RBS_LACSA | O40250 lactuca sat |
| 31 | 21 | 100.0 | 182 | 1 | PAAD_ARCFU | O29054 archaeoglob |
| 32 | 21 | 100.0 | 186 | 1 | P152_METTH | O26255 methanobact |
| 33 | 21 | 100.0 | 190 | 1 | APT_TREPA | O84001 treponema p |

| | | | | | | |
|----|----|-------|-----|---|------------|--------------------|
| 34 | 21 | 100.0 | 195 | 1 | VP12_BPPH6 | P07580 bacterioph |
| 35 | 21 | 100.0 | 200 | 1 | ACD2_CLOAB | O97dq1 clostridium |
| 36 | 21 | 100.0 | 200 | 1 | RS8A_SCHPO | O14049 schizosacch |
| 37 | 21 | 100.0 | 200 | 1 | RS8B_SCHPO | O9p7b2 schizosacch |
| 38 | 21 | 100.0 | 202 | 1 | YD99_HAETN | P44175 haemophilus |
| 39 | 21 | 100.0 | 205 | 1 | YWRF_BACST | O05220 bacillus su |
| 40 | 21 | 100.0 | 206 | 1 | YM80_YEAST | O04019 saccharomyc |
| 41 | 21 | 100.0 | 207 | 1 | GTAL_CAEEL | O09596 caenorhabdi |
| 42 | 21 | 100.0 | 208 | 1 | RS8_CAEEL | P48156 caenorhabdi |
| 43 | 21 | 100.0 | 209 | 1 | YM78_BACHD | O9rak9 bacillus ha |
| 44 | 21 | 100.0 | 210 | 1 | RS8_ARATH | O9fif3 arabidopsis |
| 45 | 21 | 100.0 | 217 | 1 | YAEF_ECOLI | P31547 escherichia |

ALIGNMENTS

RESULT 1
SUGA_ACHDO STANDARD; PRT; 5 AA.
ID SUGA_ACHDO
AC P19991;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Suboesophageal ganglion pentapeptide.
OS Acheta domestica (House cricket).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Ensifera;
OC Gryllidae; Gryllinae; Acheta.
OX NCBI_TaxID=6997;
RN [1]
RP SEQUENCE.
RA Wicker C., Wicker C.;
RT "Isolation and structure of a peptide isolated from the
RT suboesophageal ganglion of Acheta domestica (Orthoptera).";
RL Comp. Biochem. Physiol. 88C:185-187(1987).
CC -I- SUBCELLULAR LOCATION: MAIN PEPTIDE FROM THE SUBOESOPHAGEAL
CC GANGLIA.
DR PIR; JS0319; JS0319.
SQ SEQUENCE 5 AA; 476 MW; 69D76DDDDDB00000 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 2 AAPF 5

RESULT 2
COXO_BOVIN STANDARD; PRT; 63 AA.
ID COXO_BOVIN
AC P00430;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide VIIC, mitochondrial precursor
DE (EC 1.9.3.1) (VIIC).
GN COX7C OR COX7CPI.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90045968; PubMed=2554257;
RA Aqua M.S., Lomax M.I., Schon E.A., Grossman L.I.;
RT "Nucleotide sequence of a cDNA for bovine cytochrome c oxidase
RL subunit VIIC.";
RL Nucleic Acids Res. 17:8376-8376(1989).
RN [2]

```

RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=92009215; PubMed=1655579;
RA Aqua M.S., Bachman N.J., Lomax M.I., Grossman L.I.;
RT "Characterization and expression of a cDNA specifying subunit VIIC of
RL bovine cytochrome c oxidase.";
RN Gene 104:211-217(1991).
RP SEQUENCE FROM N.A.
RX MEDLINE=97248578; PubMed=9092564;
RA Seelan R.S., Grossman L.I.;
RT "Structural organization and promoter analysis of the bovine
RL cytochrome c oxidase subunit VIIC gene. A functional role for Y11.";
RN J. Biol. Chem. 272:10175-10181(1997).
RP SEQUENCE OF 17-63.
RC TISSUE=Heart;
RX MEDLINE=79046803; PubMed=213363;
RA Buse G., Steffens G.J.;
RT "Studies on cytochrome c oxidase, II. The chemical constitution of a
RL short polypeptide from the beef heart enzyme.";
RN Hoppe-Seyler's Z. Physiol. Chem. 359:1005-1009(1978).
RP SEQUENCE OF 17-42.
RC TISSUE=Liver;
RX MEDLINE=89000697; PubMed=2844245;
RA Yanamura W., Zhang Y.-Z., Takamiya S., Capaldi R.A.;
RT "Tissue-specific differences between heart and liver cytochrome c
RL oxidase.";
RN Biochemistry 27:4909-4914(1988).
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=96216288; PubMed=8638158;
RA Tsukihara T., Aoyama H., Yamashita E., Tomizaki T., Yamaguchi H.,
RT Shinzawa-Itch K., Nakashima R., Yaono R., Yoshikawa S.;
RL "The whole structure of the 13-subunit oxidized cytochrome c oxidase
RN at 2.8 A.";
RP Science 272:1136-1144(1996).
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=99190827; PubMed=10089392;
RA Tomizaki T., Yamashita E., Yamaguchi H., Aoyama H., Tsukihara T.,
RT Shinzawa-Itch K., Nakashima R., Yaono R., Yoshikawa S.;
RL "Structure analysis of bovine heart cytochrome c oxidase at 2.8 A
RN resolution.";
RP Acta Crystallogr. D 55:31-45(1999).
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE=20235645; PubMed=10771420;
RA Fei M.J., Yamashita E., Inoue N., Yao M., Yamaguchi H., Tsukihara T.,
RT Shinzawa-Itch K., Nakashima R., Yoshikawa S.;
RL "X-ray structure of azide-bound fully oxidized cytochrome c oxidase
RN from bovine heart at 2.9 A resolution.";
RP Acta Crystallogr. D 56:529-535(2000).
RP FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
RN CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
RP MITOCHONDRIAL ELECTRON TRANSPORT.
RP CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
RN c + 2 H(2)O.
RP SUBCELLULAR LOCATION: Mitochondrial inner membrane.
RP TISSUE SPECIFICITY: LIVER, HEART, MUSCLE AND BRAIN, CONTAIN THE
RN SAME ISOFORM OF COX VIIC, BUT AT DIFFERENT CONCENTRATIONS.
RP SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIC FAMILY.
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RP modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.isb-sib.ch/announce/
RP or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: X15725; CAB57793.1; -
DR EMBL: X58823; CAA41627.1; -
DR EMBL: U58655; AAC48719.1; -
DR PIR: JH0473; OSBO8A.
DR PDB: 1OCC; 07-DEC-96.
DR PDB: 2OCC; 13-JAN-99.
DR PDB: 1OCR; 29-JUL-99.
DR PDB: 1OCO; 22-JUL-99.
DR PDB: 1OCZ; 22-JUL-99.
DR InterPro: IPR004202; COX7C.
DR Pfam: PF02935; COX7C; 1.
KW Oxidoreductase; Mitochondrion; Transit peptide; 3D-structure;
KW Inner membrane; Transmembrane.
FT TRANSIT 1 16 MITOCHONDRION.
FT CHAIN 17 63 CYTOCHROME C OXIDASE POLYPEPTIDE VIIC.
FT DOMAIN 17 33 MITOCHONDRIAL MATRIX.
FT TRANSMEM 34 60
FT DOMAIN 61 63
SQ SEQUENCE 63 AA; 7331 MW; C6A2AD5CC4E63C0B CRC64;
Query Match 100.0%; Score 21; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPF 4
Db 50 AAPF 53

```

```

RESULT 3
COXO_MOUSE STANDARD; PRT; 63 AA.
ID COXO_MOUSE
AC P17665;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide VIIC, mitochondrial precursor
DE (EC 1.9.3.1).
GN COX7C OR COX7C1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Heart;
RX MEDLINE=90301494; PubMed=2163523;
RA Akamatsu M., Grossman L.I.;
RT "Nucleotide sequence of a cDNA for mouse cytochrome c oxidase subunit
RN VIIC.";
RL Nucleic Acids Res. 18:3645-3645(1990).
RP FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
RN CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
RP MITOCHONDRIAL ELECTRON TRANSPORT.
RP CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
RN c + 2 H(2)O.
RP SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIC FAMILY.
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RP modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.isb-sib.ch/announce/
RP or send an email to license@isb-sib.ch).
DR EMBL: X52940; CAA37115.1; -
DR PIR: S10303; S10303.
DR MGI: 103226; COX7C.
DR InterPro: IPR004202; COX7C.
DR Pfam: PF02935; COX7C; 1.
KW Oxidoreductase; Mitochondrion; Transit peptide.

```

```

FT TRANSIT 1 16 MITOCHONDRION.
FT CHAIN 17 63 CYTOCHROME C OXIDASE POLYPEPTIDE VIIC.
SQ SEQUENCE 63 AA; 7333 MW; C6A2AD4ABFEBC7B CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 63;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
    ||||
Db 50 AAPF 53

RESULT 4
MS2B_DROME
ID MS2B_DROME STANDARD; PRT; 90 AA.
AC P33738;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Accessory gland-specific peptide 26ab precursor (Male accessory
DE gland secretory protein 355B).
GN ACP26AB OR MST26AB OR MST355B.
OS Drosophila mauritiana (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7226;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Robertson C340;
RX MEDLINE=93106377; PubMed=1361475;
RA Aguade M., Miyashita N., Langley C.H.;
RT "Polymorphism and divergence in the Mst26A male accessory gland gene
RL region in Drosophila.";
RL Genetics 132:755-770(1992).
CC -!- FUNCTION: THIS PROTEIN IS TRANSFERRED FROM MALE TO FEMALE DURING
CC MATING AND MAY AFFECT EGG-LAYING AND BEHAVIOR AFTER MATING.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLANDS OF MALES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X70898; CAA50253.1; -.
DR PIR; S30428; S30428.
DR FlyBase; FBgn0012495; Dman\Acp26Ab.
KW Behavior; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 90 ACCESSORY GLAND-SPECIFIC PEPTIDE 26AB.
SQ SEQUENCE 90 AA; 10255 MW; 1BCFC4EB69BA0B4F CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 90;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
    ||||
Db 21 AAPF 24

RESULT 5
MS2B_DROME STANDARD; PRT; 90 AA.
ID MS2B_DROME STANDARD; PRT; 90 AA.
AC P10334; 076300; 076302; 077456; Q9V432;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Accessory gland-specific peptide 26ab precursor (Male accessory
DE gland secretory protein 355B).
GN ACP26AB OR MST26AB OR MST355B OR CG9024.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=89053045; PubMed=3142802;
RA Monsma S.A., Wolfner M.F.;
RT "Structure and expression of a Drosophila male accessory gland gene
RT whose product resembles a peptide pheromone precursor.";
RL Genes Dev. 2:1063-1073(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NC1, NC2, NC3, NC4, NC5, NC6, NC7, NC8, NC9, and NC10;
RX MEDLINE=93106377; PubMed=1361475;
RA Aguade M., Miyashita N., Langley C.H.;
RT "Polymorphism and divergence in the Mst26A male accessory gland gene
RT region in Drosophila.";
RL Genetics 132:755-770(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Various strains;
RX MEDLINE=98384843; PubMed=9718731;
RA Tsaur S.-C., Ting C.-T., Wu C.-I.;
RT "Positive selection driving the evolution of a gene of male
RT reproduction, Acp26Aa, of Drosophila: II. Divergence versus
RT polymorphism.";
RL Mol. Biol. Evol. 15:1040-1046(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=MO36A, MO37A, MO47A, MO79B, MO80B, LA25, LA46, and LA108;
RX MEDLINE=99016087; PubMed=9799260;
RA Aguade M.;
RT "Different forces drive the evolution of the Acp26Aa and Acp26Ab
RT accessory gland genes in the Drosophila melanogaster species
RT complex.";
RL Genetics 150:1079-1089(1998).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazet R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis R.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
```

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrlskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: THIS PROTEIN IS TRANSFERRED FROM MALE TO FEMALE DURING
CC MATING AND MAY AFFECT EGG-LAYING AND BEHAVIOR AFTER MATING.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLANDS OF MALES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y00219; CAA68367.1; -;
DR EMBL; X70888; CAA50233.1; -;
DR EMBL; X70889; CAA50235.1; -;
DR EMBL; X70890; CAA50237.1; -;
DR EMBL; X70891; CAA50239.1; -;
DR EMBL; X70892; CAA50241.1; -;
DR EMBL; X70893; CAA50243.1; -;
DR EMBL; X70894; CAA50245.1; -;
DR EMBL; X70895; CAA50247.1; -;
DR EMBL; X70896; CAA50249.1; -;
DR EMBL; X70897; CAA50251.1; -;
DR EMBL; AF052470; AAC27996.1; -;
DR EMBL; AF052471; AAC27998.1; -;
DR EMBL; AF052472; AAC28000.1; -;
DR EMBL; AF052473; AAC28002.1; -;
DR EMBL; AF052474; AAC28004.1; -;
DR EMBL; AF052475; AAC28006.1; -;
DR EMBL; AF052476; AAC28008.1; -;
DR EMBL; AF052477; AAC28010.1; -;
DR EMBL; AF052478; AAC28012.1; -;
DR EMBL; AF052479; AAC28014.1; -;
DR EMBL; AF052480; AAC28016.1; -;
DR EMBL; AF052481; AAC28018.1; -;
DR EMBL; AF053250; AAC28791.1; -;
DR EMBL; AF053251; AAC28793.1; -;
DR EMBL; AF053252; AAC28795.1; -;
DR EMBL; AF053253; AAC28797.1; -;
DR EMBL; AF053254; AAC28799.1; -;
DR EMBL; AF053255; AAC28801.1; -;
DR EMBL; AF053256; AAC28803.1; -;
DR EMBL; AF053257; AAC28805.1; -;
DR EMBL; AF053258; AAC28807.1; -;
DR EMBL; AF053259; AAC28809.1; -;
DR EMBL; AF053260; AAC28811.1; -;
DR EMBL; AF053261; AAC28813.1; -;
DR EMBL; AF053262; AAC28815.1; -;
DR EMBL; AF053263; AAC28817.1; -;
DR EMBL; AF053264; AAC28819.1; -;
DR EMBL; AF053265; AAC28821.1; -;
DR EMBL; AF053266; AAC28823.1; -;
DR EMBL; AF053267; AAC28825.1; -;
DR EMBL; AF053268; AAC28827.1; -;
DR EMBL; AF053269; AAC28829.1; -;
DR EMBL; AF053270; AAC28831.1; -;
DR EMBL; AF053271; AAC28833.1; -;
DR EMBL; AF053272; AAC28835.1; -;
DR EMBL; AF053273; AAC28837.1; -;
DR EMBL; AF053274; AAC28839.1; -;
DR EMBL; AF053275; AAC28841.1; -;

DR EMBL; AF053276; AAC28843.1; -;
DR EMBL; AJ231353; CAB37201.1; -;
DR EMBL; AJ231354; CAB37203.1; -;
DR EMBL; AJ231356; CAB37207.1; -;
DR EMBL; AJ231358; CAB37211.1; -;
DR EMBL; AJ231359; CAB37213.1; -;
DR EMBL; AJ231365; CAB37225.1; -;
DR EMBL; AJ231372; CAB37239.1; -;
DR EMBL; AJ231379; CAB37251.1; -;
DR EMBL; AE003611; AAF52297.1; -;
DR PIR; S02854; S02854.
DR PIR; S30408; S30408.
DR PIR; S30410; S30410.
DR PIR; S30424; S30424.
DR FlyBase; FBgn0002856; Acp26Ab.
KW Behavior; Signal; Polymorphism.
FT SIGNAL 1 21
FT CHAIN 22 90
FT VARIANT 64 64
FT VARIANT 77 77
FT VARIANT 80 80
FT VARIANT 89 89
FT VARIANT 90 90
SQ SEQUENCE 90 AA; 10162 MW; 1FDFDC7951F9BDF CRC64;

Query Match 100.0%; Score 21; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
Db 21 AAPF 24

RESULT 6
MS2B_DROSE STANDARD; PRT; 90 AA.
AC P33739;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Accessory gland-specific peptide 26Ab precursor (Male accessory
DE gland secretory protein 355B).
GN ACP26AB OR MST26AB OR MST355B.
OS *Drosophila sechellia* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7238;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93106377; PubMed=1361475;
RA Aguade M., Miyashita N., Langley C.H.;
RT "Polymorphism and divergence in the *Mst26A* male accessory gland gene
RT region in *Drosophila*.";
RL Genetics 132:755-770(1992).
CC -1- FUNCTION: THIS PROTEIN IS TRANSFERRED FROM MALE TO FEMALE DURING
CC MATING AND MAY AFFECT EGG-LAYING AND BEHAVIOR AFTER MATING.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLANDS OF MALES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial


```

CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: X72630; CAA51207.1;
DR  FlyBase; FBgn0012778; Dsec\Acp26ab.
KW  Behavior; Signal.
FT  SIGNAL 1 21
FT  CHAIN 22 90
SQ  SEQUENCE 90 AA; 10270 MW; 99CFC4EE69B7782E CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 90;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   |||
Db 21 AAPF 24

RESULT 7
MS2B_DROSI
ID MS2B_DROSI STANDARD; PRT; 90 AA.
AC P33740;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Accessory gland-specific peptide 26ab precursor (Male accessory
DE gland secretory protein 355B).
GN ACP26AB OR MST26AB OR MSF355B.
OS Drosophila simulans (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7240;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PC;
RX MEDLINE=93106377; Pubmed=1361475;
RA Aguade M., Miyashita N., Langley C.H.;
RT "Polymorphism and divergence in the Mst26A male accessory gland gene
RT region in Drosophila."
RL Genetics 132:755-770(1992).
CC -!- FUNCTION: THIS PROTEIN IS TRANSFERRED FROM MALE TO FEMALE DURING
CC MATING AND MAY AFFECT EGGLAYING AND BEHAVIOR AFTER MATING.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLANDS OF MALES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X70899; CAA50255.1;
DR PIR: S30430; S30430.
DR FlyBase; FBgn0012821; Dsim\Acp26ab.
KW Behavior; Signal.
FT SIGNAL 1 21
FT CHAIN 22 90
SQ SEQUENCE 90 AA; 10189 MW; F7F8D6E6D04F03E5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 90;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   |||
Db 21 AAPF 24

RESULT 8
U185_ORYSA
ID U185_ORYSA STANDARD; PRT; 102 AA.
AC O94DM8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein P0483G10.15.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0483G10."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE UPF0185 (BM-002) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP003263; BAB63621.1;
DR InterPro: IPR005375; UPF0185.
DR Pfam: PF03671; UPF0185; 1.
KW Hypothetical protein.
SQ SEQUENCE 102 AA; 10362 MW; 1E8F85B27F371159 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 102;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   |||
Db 99 AAPF 102

RESULT 9
YD86_MYCTU
ID YD86_MYCTU STANDARD; PRT; 102 AA.
AC P71656;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV1386 precursor.
GN RV1386 OR MT1430 OR MTCY21B4.03.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; Pubmed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
[2]

```

```
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Winn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z80108; CAB02191.1; -.
DR EMBL; AE007014; AAK45695.1; -.
DR TIGR; MT1430; -.
DR Tuberculist; RV1386; -.
DR InterPro: IPR000084; PE_region.
DR Pfam; PF00934; PE; 1.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 102 HYPOTHETICAL PROTEIN RV1386.
SQ SEQUENCE 102 AA; 9862 MW; 8165F09DBDB9D752 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 31 AAPF 34

RESULT 10
CAL2_HUMAN
ID CAL2_HUMAN STANDARD; PRT; 127 AA.
AC P10092; Q9UCN9;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcitonin gene-related peptide II precursor (CGRP-II) (Beta-type
DE CGRP).
GN CALCB OR CALC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87105923; Pubmed=3492393;
RA Steenbergh P.H., Hoeppener J.W.M., Zandberg J., Visser A.,
RA Lips C.J.M., Jansz H.S.;
RT "Structure and expression of the human calcitonin/CGRP genes.";
RL FEBS Lett. 209:97-103(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 56-127 FROM N.A.
RX MEDLINE=85180007; Pubmed=2985435;
RA Steenbergh P.H., Hoeppener J.W.M., Zandberg J., Lips C.J.M.,
RA Jansz H.S.;
RT "A second human calcitonin/CGRP gene.";
RL FEBS Lett. 183:403-407(1985).
```

```
RN [4]
RP PARTIAL SEQUENCE OF 82-108.
RX MEDLINE=87109142; Pubmed=3492492;
RA Petermann J.B., Born W., Chang J.Y., Fischer J.A.;
RT "Identification in the human central nervous system, pituitary, and
RT thyroid of a novel calcitonin gene-related peptide, and partial amino
RT acid sequence in the spinal cord.";
RL J. Biol. Chem. 262:542-545(1987).
RN [5]
RP SEQUENCE OF 82-86 AND 104-117.
RC TISSUE=Spinal cord;
RX MEDLINE=90211348; Pubmed=2322288;
RA Wimalawansa S.J., Morris H.R., Etienne A., Blench I., Panico M.,
RA McIntyre I.;
RT "Isolation, purification and characterization of beta-hCGRP from
RT human spinal cord.";
RL Biochem. Biophys. Res. Commun. 167:993-1000(1990).
RN [6]
RP SEQUENCE OF 82-104.
RC TISSUE=pheochromocytoma;
RX MEDLINE=92287083; Pubmed=1318039;
RA Kitamura K., Kangawa K., Kawamoto M., Ichiki Y., Matsuo H., Eto T.;
RT "Isolation and characterization of peptides which act on rat
RT platelets, from a pheochromocytoma.";
RL Biochem. Biophys. Res. Commun. 185:134-141(1992).
CC -1- FUNCTION: CGRP induces vasodilatation. It dilates a variety of
CC vessels including the coronary, cerebral and systemic vasculature.
CC Its abundance in the CNS also points toward a neurotransmitter or
CC neuromodulator role.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04855; CAC05295.1; -.
DR EMBL; X04857; CAC05295.1; JOINED.
DR EMBL; X04861; CAC05295.1; JOINED.
DR EMBL; BC008428; AAH08428.1; -.
DR EMBL; X02404; CAA26249.1; -.
DR PIR; A25864; A25864.
DR PIR; A34565; A34565.
DR PIR; B26142; B26142.
DR Genew; HGNC:1438; CALCB.
DR MIM; 114160; -.
DR InterPro: IPR001693; Calcitonin-like.
DR InterPro: IPR002163; Calcitonin_B.
DR Pfam; PF00214; Calc_CGRP_IAPP; 1.
DR PRINTS; PR00817; CALCITONINB.
DR SMART; SM00113; CALCITONIN; 1.
DR PROSITE; PS00258; CALCITONIN; 1.
KW Cleavage on pair of basic residues; Amidation; Hormone; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 79
FT PEPTIDE 82 118 CALCITONIN GENE-RELATED PEPTIDE II.
FT PROPEP 124 127
FT DISULFID 83 88
FT MOD_RES 118 118 AMIDATION (G-119 PROVIDE AMIDE GROUP).
FT CONFLICT 73 73 G -> S (IN REF. 3).
SQ SEQUENCE 127 AA; 13706 MW; B0A71A063CD5ACE7 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 25 AAPF 28
```

```
RESULT 11
Y355_TREPA STANDARD; PRT; 127 AA.
ID Y355_TREPA
AC 083374;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0355.
GN TP0355.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Ullnerback T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001215; AAC65353.1; -.
DR TIGR; TP0355; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 127 AA; 14526 MW; 1B848EF0DAC4ADC4 CRC64;

Query Match          100.0%; Score 21; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   1111
Db 85 AAPF 88

RESULT 12
CAL1_HUMAN STANDARD; PRT; 128 AA.
ID CAL1_HUMAN
AC P06881; Q93048; Q9UCP0;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcitonin gene-related peptide I precursor (CGRP-I) (Alpha-type
DE CGRP).
GN CALCA OR CALCL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89386053; PubMed=2571128;
RA Broad P.M., Symes A.J., Thakker R.V., Craig R.K.;
RT "Structure and methylation of the human calcitonin/alpha-CGRP gene.";
RL Nucleic Acids Res. 17:6999-7011(1989).
RN [2]
RP SEQUENCE FROM N.A.

RX MEDLINE=85166259; PubMed=3872459;
RA Jonas V., Lin C.R., Kawashima E., Semon D., Swanson L.W.,
RA Mermod J.-J., Evans R.M., Rosenfeld M.G.;
RT "Alternative RNA processing events in human calcitonin/calcitonin
RT gene-related peptide gene expression."
RL Proc. Natl. Acad. Sci. U.S.A. 82:1994-1998(1985).
RN [3]
RP SEQUENCE OF 48-119 FROM N.A.
RX MEDLINE=85022523; PubMed=6148938;
RA Nelkin B.D., Rosenfeld K.I., de Bustros A., Leong S.S., Roos B.A.,
RA Baylin S.B.;
RT "Structure and expression of a gene encoding human calcitonin and
RT calcitonin gene related peptide."
RL Biochem. Biophys. Res. Commun. 123:648-655(1984).
RN [4]
RP SEQUENCE OF 49-128 FROM N.A.
RX MEDLINE=85230541; PubMed=2408883;
RA Edbrooke M.R., Parker D., Mcvey J.H., Riley J.H., Sorenson G.D.,
RA Pettengill O.S., Craig R.K.;
RT "Expression of the human calcitonin/CGRP gene in lung and thyroid
RT carcinoma."
RL EMBO J. 4:715-724(1985).
RN [5]
RP SEQUENCE OF 77-128 FROM N.A.
RC TISSUE=Thyroid carcinoma;
RX MEDLINE=84240176; PubMed=6610687;
RA Steenbergh P.H., Hoppenner J.W., Zandberg J., de Ven W.J., Jansz H.S.,
RA Lips C.J.;
RT "Calcitonin gene related peptide coding sequence is conserved in the
RT human genome and is expressed in medullary thyroid carcinoma."
RL J. Clin. Endocrinol. Metab. 59:358-360(1984).
RN [6]
RP SEQUENCE OF 77-128 FROM N.A.
RX MEDLINE=87213363; PubMed=3034287;
RA Craig R.K., Riley J.H., Edbrooke M.R., Broad P.M., Foord S.M.,
RA Al-Kazwini S.J., Holman J.J., Marshall I.;
RT "Expression and function of the human calcitonin/alpha-CGRP gene in
RT health and disease."
RL Biochem. Soc. Symp. 52:91-105(1986).
RN [7]
RP SEQUENCE OF 83-119.
RX MEDLINE=84191466; PubMed=6609312;
RA Morris H.R., Panico M., Etienne T., Tipples J., Girgis S.I.,
RA McIntyre I.;
RT "Isolation and characterization of human calcitonin gene-related
RT peptide."
RL Nature 308:746-748(1984).
RN [8]
RP PARTIAL SEQUENCE OF 83-117.
RX MEDLINE=87109142; PubMed=3492492;
RA Petermann J.B., Born W., Chang J.Y., Fischer J.A.;
RT "Identification in the human central nervous system, pituitary, and
RT thyroid of a novel calcitonin gene-related peptide, and partial amino
RT acid sequence in the spinal cord."
RL J. Biol. Chem. 262:542-545(1987).
RN [9]
RP SEQUENCE OF 83-108, AND FUNCTION.
RC TISSUE=Pheochromocytoma;
RX MEDLINE=92287083; PubMed=1318039;
RA Kitamura K., Kangawa K., Kawamoto M., Ichiki Y., Matsuo H., Eto T.;
RT "Isolation and characterization of peptides which act on rat
RT platelets, from a pheochromocytoma."
RL Biochem. Biophys. Res. Commun. 185:134-141(1992).
RN [10]
RP STRUCTURE BY NMR OF CGRP.
RX MEDLINE=91105142; PubMed=1988044;
RA Breeze A.L., Harvey T.S., Bazzo R., Campbell I.D.;
RT "Solution structure of human calcitonin gene-related peptide by 1H
RT NMR and distance geometry with restrained molecular dynamics."
RL Biochemistry 30:575-582(1991).
RN [11]
RP STRUCTURE BY NMR OF CGRP.
RX MEDLINE=91248117; PubMed=2039456;
```

```
RA Hubbard J.A.M., Martin S.R., Chaplin L.C., Bose C., Kelly S.M.,
RA Price N.C.;
RT "Solution structures of calcitonin-gene-related-peptide analogues of
RT calcitonin-gene-related peptide and amylin.";
RL Blochem. J. 275:785-788(1991).
CC -1- FUNCTION: CGRP induces vasodilatation. It dilates a variety of
CC vessels including the coronary, cerebral and systemic vasculature.
CC Its abundance in the CNS also points toward a neurotransmitter or
CC neuromodulator role. It also elevates platelet cAMP.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (AC P01258), 2 (AC P01258) AND
CC 3 (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15943; CAA34070.1; -
DR EMBL; X02330; CAA26190.1; -
DR EMBL; M12667; AAA51914.1; -
DR EMBL; M12664; AAA51914.1; JOINED.
DR EMBL; M12665; AAA51914.1; JOINED.
DR EMBL; K03512; AAA52011.1; -
DR EMBL; M28637; AAA52012.1; -
DR EMBL; M26094; AAA51912.1; -
DR PIR; B22716; TCHUR.
DR PIR; A22949; A22949.
DR PIR; A26142; A26142.
DR PIR; S07644; S07644.
DR Genew; HGNC:1437; CALCA.
DR MIM; 114130; -
DR InterPro; IPR001693; Calcitonin-like.
DR InterPro; IPR002163; Calcitonin_B.
DR Pfam; PF00214; Calc_CGRP_IAPP; 1.
DR PRINTS; PR00817; CALCITONINB.
DR SMART; SM00113; CALCITONIN; 1.
DR PROSITE; PS00258; CALCITONIN; 1.
KW Cleavage on pair of basic residues; Amidation; Alternative splicing;
KW Hormone; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 80
FT PEPTIDE 83 119 CALCITONIN GENE-RELATED PEPTIDE I.
FT PROPEP 125 128
FT DISULFID 84 89
FT MOD_RES 119 119 AMIDATION (G-120 PROVIDE AMIDE GROUP).
FT CONFLICT 48 48 R -> V (IN REF. 4).
FT CONFLICT 76 76 R -> V (IN REF. 3).
SQ SEQUENCE 128 AA; 13899 MW; A003A1069260D9B8 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
ID 1111
DB 25 AAPF 28

RESULT 13
CAL2_MOUSE STANDARD; PRT; 130 AA.
AC Q99MP3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcitonin gene-related peptide II precursor (CGRP-II) (Beta-type
DE CGRP).
GN CALCB.
OS Mus musculus (Mouse).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=21604266; PubMed=11761712;
RA Thomas P.M., Nasonkin I., Zhang H., Gagel R.F., Cote G.J.;
RT "Structure of the mouse calcitonin/calcitonin gene-related peptide
RT alpha and beta genes.";
RL DNA Seq. 12:131-135(2001).
CC -1- FUNCTION: CGRP induces vasodilatation. It dilates a variety of
CC vessels including the coronary, cerebral and systemic vasculature.
CC Its abundance in the CNS also points toward a neurotransmitter or
CC neuromodulator role (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF325526; AAK16431.1; -
DR EMBL; AF325524; AAK16431.1; JOINED.
DR MGD; MGI:2151254; Calcb.
DR InterPro; IPR000443; Amylin.
DR InterPro; IPR001693; Calcitonin-like.
DR InterPro; IPR002163; Calcitonin_B.
DR Pfam; PF00214; Calc_CGRP_IAPP; 1.
DR PRINTS; PR00817; CALCITONINB.
DR PRINTS; PR00818; ISLETAMYLID.
DR SMART; SM00113; CALCITONIN; 1.
DR PROSITE; PS00258; CALCITONIN; 1.
KW Cleavage on pair of basic residues; Amidation; Hormone; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT PROPEP 27 82 BY SIMILARITY.
FT PEPTIDE 84 120 CALCITONIN GENE-RELATED PEPTIDE II.
FT PROPEP 127 130 BY SIMILARITY.
FT DISULFID 85 90 BY SIMILARITY.
FT MOD_RES 120 120 AMIDATION (G-121 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
SQ SEQUENCE 130 AA; 14623 MW; 97299244E8F6C536 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
ID 1111
DB 26 AAPF 29

RESULT 14
SZ05_RAT STANDARD; PRT; 130 AA.
AC P97885;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Small inducible cytokine B5 precursor (CXCL5) (Cytokine LIX).
DE SCYB5.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Kelnner G.S., Maciejewski-Lenoir D., Lee E.D., Maki R.A.;
```


RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY PARTICIPATE IN THE RECRUITMENT OF INFLAMMATORY CELLS
CC BY INJURED OR INFECTED TISSUE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC C-X-C) (CHEMOKINE CXC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U90448; AAB61460.1; -.
DR HSSP; P10889; IMI2.
DR InterPro; IPR001089; CXC_chmkine_sml.
DR InterPro; IPR001811; Chemokine_IL8.
DR Pfam; PF00048; IL8; 1.
DR PRINTS; PR00437; SMALLCYTKCXC.
DR SMART; SM00199; SCY; 1.
DR PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
KW Cytokine; Signal.
FT SIGNAL 1 37 POTENTIAL.
FT CHAIN 38 130 SMALL INDUCIBLE CYTOKINE B5.
FT DISULFID 50 76 BY SIMILARITY.
FT DISULFID 52 93 BY SIMILARITY.
SQ SEQUENCE 130 AA; 14263 MW; C00F6B3605524F4E CRC64;

Query Match 100.0%; Score 21; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 37 AAPF 40

RESULT 15
P2Y2_BOVIN STANDARD; PRT; 133 AA.
ID P2Y2_BOVIN
AC O18951;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE P2Y purinoceptor 2 (P2Y2) (Fragment).
GN P2RY2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Chang A.S., Chang S.M., Schilling W.P.;
RT "Identification of multiple p2 purinergic receptor subtypes in bovine
RT aortic endothelial cells (BAECs).";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
CC activate a phosphatidylinositol-calcium second messenger system
CC (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF005153; AAB62543.1; -.

DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECIP_F1_1; FALSE_NEG.
DR PROSITE; PS50262; G_PROTEIN_RECIP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane.
FT NON_TER 1 1
FT DOMAIN <1 26 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 27 47 4 (POTENTIAL).
FT DOMAIN 48 74 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 75 95 5 (POTENTIAL).
FT DOMAIN 96 116 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 117 133 6 (POTENTIAL).
FT NON_TER 133 133
SQ SEQUENCE 133 AA; 15006 MW; 301B087C62C75B69 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 82 AAPF 85

RESULT 16
RBFA_SYNY3 STANDARD; PRT; 133 AA.
ID RBFA_SYNY3
AC Q55625;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribosome-binding factor A.
GN RBFA OR SL0754.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- FUNCTION: ASSOCIATES WITH FREE 30S RIBOSOMAL SUBUNITS (BUT NOT
CC WITH 30S SUBUNITS THAT ARE PART OF 70S RIBOSOMES OR POLYSOMES).
CC ESSENTIAL FOR EFFICIENT PROCESSING OF 16S RNA. MAY INTERACT WITH
CC THE 5'TERMINAL HELIX REGION OF 16S RNA (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE RBFA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D64000; BAA10137.1; -.
DR InterPro; IPR000238; Rfb_bind_facta.
DR Pfam; PF02033; RBFA; 1.
DR ProDom; PD007327; Rfb_bind_facta; 1.
DR TIGRFAMs; TIGR00082; rbfA; 1.
DR PROSITE; PS01319; RBFA; 1.
KW rRNA processing; Complete proteome.
SQ SEQUENCE 133 AA; 14749 MW; 1A7C076E9F7B9829 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
 Db 73 AAPF 76

```

RESULT 17
CAL2_RAT STANDARD; PRT; 134 AA.
AC P10093;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcitonin gene-related peptide II precursor (CGRP-II) (Beta-type
DE CGRP).
DE CALCB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85300490; PubMed=2994212;
RA Amara S.G., Arriaza J.L., Lefk S.E., Swanson L.W., Evans R.M.,
RA Rosenfeld M.G.;
RT "Expression in brain of a messenger RNA encoding a novel neuropeptide
RT homologous to calcitonin gene-related peptide."
RL Science 229:1094-1097(1985).
CC -1- FUNCTION: CGRP induces vasodilatation. It dilates a variety of
CC vessels including the coronary, cerebral and systemic vasculature.
CC Its abundance in the CNS also points toward a neurotransmitter or
CC neuromodulator role.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M1596; AAA40850.1; -.
DR PIR; A44173; A44173.
DR InterPro; IPR001693; Calcitonin-like.
DR InterPro; IPR002163; Calcitonin_B.
DR Pfam; PF00214; Calc_CGRP_IAPP; 1.
DR PRINTS; PR00817; CALCITONINB.
DR SMART; SM00113; CALCITONIN; 1.
DR PROSITE; PS00258; CALCITONIN; 1.
KW Cleavage on pair of basic residues; Amidation; Hormone; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT PROPEP 27 86 BY SIMILARITY.
FT PEPTIDE 89 125 CALCITONIN GENE-RELATED PEPTIDE II.
FT PROPEP 131 134 BY SIMILARITY.
FT DISULFID 90 95 BY SIMILARITY.
FT MOD_RES 125 125 AMIDATION (G-126 PROVIDE AMIDE GROUP) (BY
FT SEQUENCE 134 AA; 14965 MW; BF6CAFB87A489B38 CRC64;
  SO
  Query Match 100.0%; Score 21; DB 1; Length 134;
  Best Local Similarity 100.0%; Pred. No. 1.7e+02;
  Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 1 AAPF 4
  Db 26 AAPF 29
  RESULT 18,
  CU3A_TENMO STANDARD; PRT; 134 AA.
  ID CU3A_TENMO
  AC P80683;

```

```

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Larval cuticle protein A3A (TM-A3A) (TM-LCP A3A).
OS Tenebrio molitor (Yellow mealworm).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Tenebrionidae; Tenebrio.
OX NCBI_TaxID=7067;
RN [1]
RP SEQUENCE.
RC TISSUE=Cuticle;
RX MEDLINE=97218698; PubMed=9066122;
RA Andersen S.O., Rafn K., Roepstorff P.;
RT "Sequence studies of proteins from larval and pupal cuticle of the
RT yellow meal worm, Tenebrio molitor."
RL Insect Biochem. Mol. Biol. 27:121-131(1997).
CC -1- FUNCTION: COMPONENT OF THE CUTICLE OF THE LARVA OF TENEBRIO
CC MOLITOR.
CC -1- DOMAIN: THE TETRAPEPTIDE (A-A-P-[AV]) REPEATS FOUND THROUGHOUT THE
CC PROTEIN ARE ALSO PRESENT IN MANY PROTEINS CONSTITUTING THE
CC PROTECTIVE ENVELOPE OF OTHER SPECIES.
CC -1- MASS SPECTROMETRY: MW=13974.3; METHOD=Electrospray.
CC -1- SIMILARITY: CONTAINS 1 CUTICLE CONSENSUS DOMAIN.
DR InterPro; IPR000618; Insect_cuticle.
DR Pfam; PF00379; Chitin_bind_4; 1.
DR PRINTS; PR00947; CUTICLE.
DR PROSITE; PS00233; CUTICLE; 1.
KW Structural protein; Cuticle; Repeat.
FT REPEAT 23 26 1.
FT REPEAT 111 114 2.
FT SEQUENCE 134 AA; 13976 MW; F6AAAF06444FE93E CRC64;
  SO
  Query Match 100.0%; Score 21; DB 1; Length 134;
  Best Local Similarity 100.0%; Pred. No. 1.7e+02;
  Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 1 AAPF 4
  Db 14 AAPF 17

```

```

RESULT 19
NU2M_ARTSA STANDARD; PRT; 136 AA.
ID NU2M_ARTSA
AC P19042;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 2 (EC 1.6.5.3) (Fragments).
GN ND2.
OS Artemia salina (Brine shrimp).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Branchiopoda; Anostraca; Artemiidae; Artemia.
OX NCBI_TaxID=85549;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88289417; PubMed=3135541;
RA Batuecas B., Garesse R., Calleja M., Valverde J.R., Marco R.;
RT "Genome organization of Artemia mitochondrial DNA.";
RL Nucleic Acids Res. 16:6515-6529(1988).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone -> NAD(+) + ubiquinol.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC      or send an email to license@lsb-sib.ch).
CC      -----
CC      EMBL; X07670; CAA30517.1; -.
DR      EMBL; X07671; CAA30518.1; -.
KW      PIR; S01220; S01220.
KW      Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
FT      NON_TER      1      1
FT      NON_CONS      83      84
FT      NON_TER      136      136
SQ      SEQUENCE      136 AA; 15330 MW; FFD3D8FB4DE9BA17 CRC64;

Query Match      100.0%; Score 21; DB 1; Length 136;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
        ||||
Db       46 AAPF 49

RESULT 20
CALO_HUMAN
ID      CALO_HUMAN      STANDARD;      PRT;      141 AA.
AC      P01258; Q13937;
DT      21-JUL-1986 (Rel. 01, Created)
DT      01-APR-1988 (Rel. 07, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Calcitonin precursor [Contains: Calcitonin; Katalcalcin (Calcitonin
DE      carboxyl-terminal peptide) (CCP) (PDN-21)].
OS      CALCA OR CALCI.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RX      MEDLINE=89386053; Pubmed=2571128;
RA      Broad P.M., Symes A.J., Thakker R.V., Craig R.K.;
RT      "Structure and methylation of the human calcitonin/alpha-CGRP gene.";
RL      Nucleic Acids Res. 17:6999-7011(1989).
[2]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RX      MEDLINE=84132556; Pubmed=6546550;
RA      le Moulllec J.M., Julienne A., Chenais J., Lasmoles F.,
RA      Guliana J.M., Milhaud G., Moukhtar M.S.;
RT      "The complete sequence of human preprocalcitonin.";
RL      FEBS Lett. 167:93-97(1984).
[3]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RX      MEDLINE=85166259; Pubmed=3872459;
RA      Jonas V., Lin C.R., Kawashima E., Semon D., Swanson L.W.,
RA      Mermod J.-J., Evans R.M., Rosenfeld M.G.;
RT      "Alternative RNA processing events in human calcitonin/calcitonin
RT      gene-related peptide gene expression.";
RL      Proc. Natl. Acad. Sci. U.S.A. 82:1994-1998(1985).
[4]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RX      MEDLINE=86164952; Pubmed=3485540;
RA      Riley J.H., Edbrooke M.R., Craig R.K.;
RT      "Ectopic synthesis of high-Mr calcitonin by the BEN lung carcinoma
RT      cell line reflects aberrant proteolytic processing.";
RL      FEBS Lett. 198:71-79(1986).
[5]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RX      MEDLINE=87213363; Pubmed=3034287;
RA      Craig R.K., Riley J.H., Edbrooke M.R., Broad P.M., Foord S.M.,
RA      Al-Kazwini S.J., Holman J.J., Marshall I.;
RT      "Expression and function of the human calcitonin/alpha-CGRP gene in
RT      health and disease.";
RL      Biochem. Soc. Symp. 52:91-105(1986).
[6]
RP      SEQUENCE FROM N.A. (ISOFORM 2).
RC      TISSUE=Thyroid carcinoma;

```

```

RX MEDLINE=92105127; PubMed=1761559;
RA Minvielle S., Giscard-Darteville S., Cohen R., Taboulet J., Labye F.,
RA Jullienne A., Rivaille P., Milhaud G., Moukhtar M.S., Lasmoles F.;
RT "A novel calcitonin carboxyl-terminal peptide produced in medullary
RT thyroid carcinoma by alternative RNA processing of the
RL calcitonin/calcitonin gene-related peptide gene.";
RL J. Biol. Chem. 266:24627-24631(1991).
RN [7]
RP SEQUENCE OF 49-141 FROM N.A. (ISOFORM 1).
RX MEDLINE=85230541; PubMed=2408883;
RA Edbrooke M.R., Parker D., McVey J.H., Riley J.H., Sorenson G.D.,
RT Peltengill O.S., Craig R.K.;
RT "Expression of the human calcitonin/CGRP gene in lung and thyroid
RT carcinoma.";
RL EMBO J. 4:715-724(1985).
RN [8]
RP SEQUENCE OF 48-83 FROM N.A.
RX MEDLINE=85022523; PubMed=6148938;
RA Nelkin B.D., Rosenfeld K.I., de Bustros A., Leong S.S., Roos B.A.,
RA Baylin S.B.;
RT "Structure and expression of a gene encoding human calcitonin and
RT calcitonin gene related peptide.";
RL Biochem. Biophys. Res. Commun. 123:648-655(1984).
RN [9]
RP SEQUENCE OF 85-116.
RX MEDLINE=72162720; PubMed=5760861;
RA Neher R., Riniker B., Rittel W., Zuber H.;
RT "Human calcitonin. Structure of calcitonin M and D.";
RL Helv. Chim. Acta 51:1900-1905(1968).
RN [10]
RP STRUCTURE BY NMR OF CALCITONIN.
RX MEDLINE=91159414; PubMed=2001366;
RA Motta A., Temussi P.A., Munsch E., Bovermann G.;
RT "A 1H NMR study of human calcitonin in solution.";
RL Biochemistry 30:2364-2371(1991).
RN [11]
RP CHARACTERIZATION OF KATACALCIN.
RX MEDLINE=8316029; PubMed=6132180;
RA Hillyard C.J., Myers C., Abeyasekera G., Stevensvenson J.C.,
RA Craig R.K., Macintyre I.;
RT "Katacalcic: a new plasma calcium-lowering hormone.";
RL Lancet 1:846-848(1983).
CC -!- FUNCTION: CALCITONIN CAUSES A RAPID BUT SHORT-LIVED DROP IN THE
CC LEVEL OF CALCIUM AND PHOSPHATE IN BLOOD BY PROMOTING THE
CC INCORPORATION OF THOSE IONS IN THE BONES.
CC -!- FUNCTION: Katacalcic is a potent plasma calcium-lowering peptide.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3 (AC
CC P06881); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15943; CAA34070.1; ALT_SEQ.
DR EMBL; X00356; CAA25103.1; -
DR EMBL; K03513; AAA52124.1; -
DR EMBL; M12666; AAA51913.1; -
DR EMBL; M12664; AAA51913.1; JOINED.
DR EMBL; M12665; AAA51913.1; JOINED.
DR EMBL; X03662; CAA27299.1; -
DR EMBL; M64486; AAA58403.1; -
DR EMBL; M26095; AAA35501.1; -
DR PIR; A22716; TCHU.
DR PIR; C22949; C22949.
DR PIR; S07238; S07238.
DR PIR; S07242; S07242.
DR PIR; S07643; S07643.

```

DR HSSP: P01262; 1BKU.
DR Genew; HGNC:1437; CALCA.
DR MIM; 114130; .
DR InterPro; IPR001693; Calcitonin-like.
DR InterPro; IPR001935; Calcitonin_A.
DR Pfam; PF00214; Calc_CGRP_IAPP; 1.
DR PRINTS; PR00270; CALCITONINA.
DR SMART; SM00113; CALCITONIN; 1.
DR PROSITE; PS00258; CALCITONIN; 1.
KW Cleavage on pair of basic residues; Amidation; Alternative splicing;
KW Hormone; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 82
FT PEPTIDE 85 116 CALCITONIN.
FT PEPTIDE 121 141 KATACALCIN.
FT DISULFID 85 91
FT MOD_RES 116 116
FT VARSPLIC 134 141
FT CONFLICT 92 92 M -> I (IN REF. 3).
SQ SEQUENCE 141 AA; 15467 MW; 99622305DD8B286F CRC64;

Query Match 100.0%; Score 21; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 25 AAPF 28

RESULT 21
YEF5_YEAST
ID YEF5_YEAST STANDARD; PRT; 141 AA.
AC P32616;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Hypothetical 16.5 kDa protein in GLY1-GDA1 intergenic region.
GN YEL045C OR SYGP-ORF33.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,
RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18779; AAB64997.1; .
DR PIR; S30832; S30832.
DR SGD; S0000771; YEL045C.
KW Hypothetical protein; ATP-binding; Transmembrane.
FT NP_BIND 15 22 ATP (POTENTIAL).

FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 67 87 POTENTIAL.
SQ SEQUENCE 141 AA; 16468 MW; F6604AC5343A5D5C CRC64;

Query Match 100.0%; Score 21; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 39 AAPF 42

RESULT 22
PSAH_MAIZE
ID PSAH_MAIZE STANDARD; PRT; 142 AA.
AC O65101;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Photosystem I reaction center subunit VI, chloroplast precursor
DE (PSI-H) (light-harvesting complex I 11 kDa protein).
GN PSAH.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. B73;
RA Heck D.A., Chitnis P.R.;
RT "Nucleotide sequence of cDNAs encoding the psah and psan subunits of
RT the maize photosystem I complex.";
RL (In) Plant Gene Register PGR98-107.
CC -1- FUNCTION: POSSIBLE ROLE COULD BE THE DOCKING OF THE LHC I ANTENNA
CC COMPLEX TO THE CORE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE PSAH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF052076; AAC26196.1; .
DR InterPro; IPR004928; PSI_Psah.
DR Pfam; PF03244; PSI_Psah; 1.
KW Photosynthesis; Photosystem I; Chloroplast; Transit peptide;
KW Transmembrane.
FT TRANSIT 1 48 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 49 142 PHOTOSYSTEM I REACTION CENTER SUBUNIT VI.
FT TRANSMEM 99 119 POTENTIAL.
SQ SEQUENCE 142 AA; 14929 MW; C8E9540210026B96 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 142;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 91 AAPF 94

RESULT 23
PSAH_HORVU
ID PSAH_HORVU STANDARD; PRT; 143 AA.
AC P20143;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

DE Photosystem I reaction center subunit VI, chloroplast precursor
DE (PSI-H) (Light-harvesting complex I 11 kDa protein).
GN PSAB.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=89325636; PubMed=2666162;
RA Okkels J.S., Scheller H.V., Jepsen L.B., Moeller B.L.;
RT "A cDNA clone encoding the precursor for a 10.2 kDa photosystem I
RT polypeptide of barley.";
RL FEBS Lett. 250:575-579(1989).
CC -1- FUNCTION: POSSIBLE ROLE COULD BE THE DOCKING OF THE LHC I ANTENNA
CC COMPLEX TO THE CORE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE PSAB FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X16092; CAA34218.1; -.
DR PIR: S05012; S05012.
DR InterPro: IPR004928; PSI_Psah.
DR Pfam: PF03244; PSI_Psah; 1.
KW Photosynthesis; Photosystem I; Chloroplast; Transit peptide;
KW Transmembrane.
FT TRANSIT 1 48 CHLOROPLAST.
FT CHAIN 49 143 PHOTOSYSTEM I REACTION CENTER SUBUNIT VI.
FT TRANSMEM 100 120 POTENTIAL.
SQ SEQUENCE 143 AA; 14882 MW; 16D72FEF36C9D46C2 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 143;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   ||||
Db 92 AAPF 95

RESULT 24
PSAB_SPIOL
ID PSAB_SPIOL STANDARD; PRT; 144 AA.
AC P22179;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Photosystem I reaction center subunit VI, chloroplast precursor
DE (PSI-H) (Light-harvesting complex I 11 kDa protein).
GN PSAB.
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90090689; PubMed=2688927;
RA Steppuhn J., Hermans J., Nechushtal R., Herrmann G.S., Herrmann R.G.;
RT "Nucleotide sequences of cDNA clones encoding the entire precursor
RT polypeptide for subunit VI and of the plastome-encoded gene for
RT subunit VII of the photosystem I reaction center from spinach.";
RL Curr. Genet. 16:99-108(1989).
CC -1- FUNCTION: POSSIBLE ROLE COULD BE THE DOCKING OF THE LHC I ANTENNA
CC COMPLEX TO THE CORE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE PSAB FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X16858; CAA34749.1; -.
DR PIR: S00453; S00453.
DR InterPro: IPR004928; PSI_Psah.
DR Pfam: PF03244; PSI_Psah; 1.
KW Photosynthesis; Photosystem I; Chloroplast; Transit peptide;
KW Transmembrane.
FT TRANSIT 1 49 CHLOROPLAST.
FT CHAIN 50 144 PHOTOSYSTEM I REACTION CENTER SUBUNIT VI.
FT TRANSMEM 104 122 POTENTIAL.
SQ SEQUENCE 144 AA; 15325 MW; 552BAB7E063375A5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 144;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   ||||
Db 93 AAPF 96

RESULT 25
PSAB_BRARA
ID PSAB_BRARA STANDARD; PRT; 145 AA.
AC O04006;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Photosystem I reaction center subunit VI, chloroplast precursor
DE (PSI-H) (Light-harvesting complex I 11 kDa protein).
GN PSAB.
OS Brassica rapa (Turnip).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=51350;
RN [1]
RP SEQUENCE FROM N.A.
RA Son D., Jo J.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: POSSIBLE ROLE COULD BE THE DOCKING OF THE LHC I ANTENNA
CC COMPLEX TO THE CORE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE PSAB FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U92504; AAB51159.1; -.
DR InterPro: IPR004928; PSI_Psah.
DR Pfam: PF03244; PSI_Psah; 1.
KW Photosynthesis; Photosystem I; Chloroplast; Transit peptide;
KW Transmembrane.
FT TRANSIT 1 50 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 51 145 PHOTOSYSTEM I REACTION CENTER SUBUNIT VI.
FT TRANSMEM 102 122 POTENTIAL.
SQ SEQUENCE 145 AA; 15410 MW; 9A0940A08EB4E790 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 145;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


OY 1 AAPF 4
1111
DB 94 AAPF 97

```
RESULT 26
PSH1_ARATH STANDARD; PRT; 145 AA.
ID PSH1_ARATH STANDARD; PRT; 145 AA.
AC Q9SUT7; Q42295;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Photosystem I reaction center subunit VI-1, chloroplast precursor
  (PSI-H1).
GN PSAH1 OR AT3G16140 OR MSL1.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Legen J., Misera S., Herrmann R.G., Altschmied L.;
RT "Sequences and map position of 31 Arabidopsis thaliana cDNAs encoding
  organellar polypeptides.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I.
  Sequence features of the regions of 4,504,864 bp covered by sixty P1
  and TAC clones.";
RL DNA Res. 7:131-135(2000).
RN [3]
RP SEQUENCE OF 31-127 FROM N.A.
RC STRAIN=cv. Columbia;
RA Bardet C., Dabos P., Tremoureygue D., Lescure B.;
RT Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: POSSIBLE ROLE COULD BE THE DOCKING OF THE LHC I ANTENNA
  COMPLEX TO THE CORE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE PSAH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ245631; CAB52749.1; -
DR EMBL; AB012247; BAB02680.1; -
DR EMBL; Z35381; CAAB4571.1; -
DR InterPro; IPR004928; PSI_PSAH.
DR Pfam; PF03244; PSI_PSAH; 2.
KW Photosynthesis; Photosystem I; Chloroplast; Transit peptide;
KW Transmembrane.
FT TRANSIT 1 50 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 51 145 PHOTOSYSTEM I REACTION CENTER SUBUNIT VI-
FT CONFLICT 112 127 1.
FT GSLTYVSATSTGEVL -> RLFGYLCQRYLYRRSS (IN
  REF. 3).
SQ SEQUENCE 145 AA; 15216 MW; 4EF2A53CDC17072E CRC64;
```

OY 1 AAPF 4

Query Match 100.0%; Score 21; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. NO. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 94 AAPF 97
1111

```
RESULT 27
PSH2_ARATH STANDARD; PRT; 145 AA.
ID PSH2_ARATH STANDARD; PRT; 145 AA.
AC Q9SUI6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Photosystem I reaction center subunit VI-2, chloroplast precursor
  (PSI-H1).
GN PSAH2 OR AT1G52230 OR P915.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Legen J., Misera S., Herrmann R.G., Altschmied L.;
RT "Sequences and map position of 31 Arabidopsis thaliana cDNAs encoding
  organellar polypeptides.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shin P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
  thaliana.";
RL Nature 408:816-820(2000).
CC -1- FUNCTION: POSSIBLE ROLE COULD BE THE DOCKING OF THE LHC I ANTENNA
  COMPLEX TO THE CORE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE PSAH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ245632; CAB52750.1; -
DR EMBL; AC022354; AAF29410.1; -
DR InterPro; IPR004928; PSI_PSAH.
DR Pfam; PF03244; PSI_PSAH; 1.
KW Photosynthesis; Photosystem I; Chloroplast; Transit peptide;
KW Transmembrane.
FT TRANSIT 1 50 CHLOROPLAST (POTENTIAL).
FT CHAIN 51 145 PHOTOSYSTEM I REACTION CENTER SUBUNIT VI-
FT CONFLICT 112 127 2.
SQ SEQUENCE 145 AA; 15273 MW; E2884F419908CD30 CRC64;
```

Query Match

100.0%; Score 21; DB 1; Length 145;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
1111
Db 94 AAPF 97

RESULT 28
VNS3_CVPPU STANDARD; PRT; 165 AA.
ID VNS3_CVPPU
AC P09047;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE Nonstructural protein 3-1 (X2B protein).
GN NS3-1.
OS Porcine transmissible gastroenteritis coronavirus (strain Purdue).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11151;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=88078100; PubMed=2825819;
RA Rassaert D., Gelfi J., Laude H.;
RT "Enteric coronavirus JGEV: partial sequence of the genomic RNA, its
RT organization and expression.";
RL Biochimie 69:591-600(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X06371; CAA29671.1; -.
DR PIR; S01740; S01740.
DR InterPro; IPR004293; Corona_NS3b.
DR Pfam; PF03053; Corona_NS3b; 1.
KW Nonstructural protein.
SQ SEQUENCE 165 AA; 18831 MW; 34F74349D3DD3B9 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
1111
Db 64 AAPF 67

RESULT 29
RBS_SACHY STANDARD; PRT; 168 AA.
ID RBS_SACHY
AC Q41373;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribulose biphosphate carboxylase small chain, chloroplast precursor
DE (EC 4.1.1.39) (Rubisco small subunit).
GN RBCS.
OS Saccharum hybrid (Sugarcane).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Saccharum.
OX NCBI_TaxID=15819;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=CV. H32-8560;
RX MEDLINE=93222494; PubMed=8467089;
RA Tang W., Sun S.S.;

RT "Sequence of a sugarcane ribulose-1,5-bisphosphate
RT carboxylase/oxygenase small subunit gene.";
RL Plant Mol. Biol. 21:949-951(1993).
CC -i- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -i- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) = 2 3-
CC phospho-D-glycerate.
CC -i- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + O(2) =
CC 3-phospho-D-glycerate + 2-phosphoglycolate.
CC -i- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -i- SUBCELLULAR LOCATION: Chloroplast.
CC -i- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M86717; AAA33922.1; -.
DR InterPro; IPR000894; Rubisco_small.
DR Pfam; PF00101; Rubisco_small; 1.
DR PRINTS; PR00152; RUBISCO SMALL.
DR ProDom; PD000290; Rubisco_small; 1.
KW Photosynthesis; Carbon dioxide fixation; Photorespiration;
KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Transit peptide.
FT TRANSIT 1 46 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 47 168 RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL
SQ SEQUENCE 168 AA; 19036 MW; DED5D100B1C314F4 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 168;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
1111
Db 13 AAPF 16

RESULT 30
RBS_LACSA STANDARD; PRT; 181 AA.
ID RBS_LACSA
AC Q40250;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribulose biphosphate carboxylase small chain, chloroplast precursor
DE (EC 4.1.1.39) (Rubisco small subunit).
GN RBCS.
OS Lactuca sativa (Garden lettuce).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Lactuca.
OX NCBI_TaxID=4236;
RN 11
RP SEQUENCE FROM N.A.
RA Hanyu K.;
RL Submitted (JAN-1993) to the EMBL/Genbank/DBJ databases.
CC -i- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -i- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) = 2 3-
CC phospho-D-glycerate.

```

CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + O(2) =
CC 3-phospho-D-glycerate + 2-phosphoglycolate.
CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; D14001; BAA03103.1; -
CC InterPro; IPR000894; Rubisco_small.
CC Pfam; PF00101; Rubisco_small; 1.
CC PRINTS; PR00152; RUBISCO_SMALL.
CC ProDom; PD000290; Rubisco_small; 1.
CC Photosynthesis; Carbon dioxide fixation; Photorespiration;
CC Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Transil peptide.
CC TRANSIT 1 57 CHLOROPLAST (BY SIMILARITY).
CC CHAIN 58 181 RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL
CC CHAIN.
CC SEQUENCE 181 AA; 20359 MW; CEF63AD4E81F6075 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 181;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
   |||
Db 23 AAPF 26

RESULT 31
PAAD_ARCFU STANDARD; PRT; 182 AA.
ID PAAD_ARCFU
AC O29054;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable aromatic acid decarboxylase (EC 4.1.1.-).
GN AF1214.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus S.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -1- SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE /
CC PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE001020; AAB90031.1; -
CC TIGR; AF1214; -
CC InterPro; IPR003382; Flavoprotein.
CC InterPro; IPR004507; ubix.
CC Pfam; PF02441; Flavoprotein; 1.
CC TIGRFAMS; TIGR00421; ubix; 1.
CC Hypothetical protein; lyase; Decarboxylase; Complete proteome.
CC SEQUENCE 182 AA; 19968 MW; 65F015F4ED130768 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
   |||
Db 67 AAPF 70

RESULT 32
P152_METTH STANDARD; PRT; 186 AA.
ID P152_METTH
AC O26255;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein MTH152.
GN MTH152.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RA "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=20473232; PubMed=11017201;
RA Christendat D., Yee A., Dharamsi A., Kluger Y., Savchenko A.,
RA Cort J.R., Booth V., Mackereith C.D., Saridakis V., Ekkel I.,
RA Kozlov G., Maxwell K.L., Wu N., McIntosh L.P., Gehring K.,
RA Kennedy M.A., Davidson A.R., Pal E.F., Gerstein M., Edwards A.M.,
RA Arowsmith C.H.;
RA "Structural proteomics of an archaeon.";
RT Nat. Struct. Biol. 7:903-909(2000).
RL -1- COFACTOR: FMN.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE FLAVOREDODIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE000803; AAB84658.1; -
CC PDB; 1EJE; 11-OCT-00.

```

```
DR InterPro; IPR002563; Flavin_Reduct.
DR Pfam; PF01613; Flavin_Reduct; 1.
KW Flavoprotein; FMN; 3D-structure; Complete proteome.
SQ SEQUENCE 186 AA; 20323 MW; 6997E8C6D9234A89 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 186;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   ||||
Db 37 AAPF 40

RESULT 33
APT_TREPA
ID APT_TREPA STANDARD; PRT; 190 AA.
AC 084001;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).
GN APT OR TP1035.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwyn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -1- FUNCTION: CATALYZES A SALVAGE REACTION RESULTING IN THE FORMATION
CC OF AMP, THAT IS ENERGICALLY LESS COSTLY THAN DE NOVO SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -1- PATHWAY: Purine salvage.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001271; AAC26593.1; -.
DR TIGR; TP1039; -.
DR InterPro; IPR000836; PRTtransferase.
DR InterPro; IPR002375; Pr/PY_rp_transf.
DR Pfam; PF00156; Pribosyltran; 1.
DR PROSITE; PS00103; PUR_PYR_PR_TRANSFER; FALSE_NEG.
KW transferase; Glycosyltransferase; Purine salvage; Complete proteome.
SQ SEQUENCE 190 AA; 20902 MW; EC12F26DAEDC5D81 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 190;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   ||||
Db 153 AAPF 156

RESULT 35
ACD2_CLOAB
ID ACD2_CLOAB STANDARD; PRT; 200 AA.
AC 097D01;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative acyl carrier protein phosphodiesterase 2 (EC 3.1.4.14) (ACP
DE phosphodiesterase 2).
GN CAC3421.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
CC -1- FUNCTION: Converts holo-ACP to apo-ACP by hydrolytic cleavage of
CC the phosphopantetheine prosthetic group from ACP (By similarity).
CC -1- CATALYTIC ACTIVITY: Holo-[acyl-carrier protein] + H(2)O = 4,-
```

```
CC      phosphopantetheine + apo-[acyl-carrier protein].
CC      -1- SIMILARITY: BELONGS TO THE ACPD FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AE007839; AAK81351.1; -
KW      Hypothetical protein; Hydrolase; Complete proteome.
SQ      SEQUENCE 200 AA; 22618 MW; FE0C21A5B62859D2 CRC64;

Query Match      100.0%; Score 21; DB 1; Length 200;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
        ||||
Db      88 AAPF 91

RESULT 36
RS8A_SCHPO
ID  RS8A_SCHPO      STANDARD;      PRT;      200 AA.
AC  014049;
DT  15-DEC-1998 (Rel. 37, Created)
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  40S ribosomal protein S8-A.
GN  RPS8A OR RPS8 OR SPAC2C4.16C.
OS  Schizosaccharomyces pombe (Fission yeast).
OC  Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC  Schizosaccharomycetales; Schizosaccharomycetaceae;
OC  Schizosaccharomyces.
OX  NCBI_TaxID=4896;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=972;
RX  MEDLINE=21848401; Pubmed=11859360;
RA  Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA  Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA  Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA  Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA  Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA  Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA  James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA  Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA  Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA  Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA  Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA  Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA  Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA  Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA  Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA  Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA  Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA  Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA  Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA  Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA  Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA  Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA  Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA  Shpakovski G.V., Ussey D., Barrell B.G., Nurse P.;
RT  "The genome sequence of Schizosaccharomyces pombe.";
RL  Nature 415:871-880(2002).
CC  -1- MISCELLANEOUS: THERE ARE TWO GENES FOR S8 IN S.POMBE.
CC  -1- SIMILARITY: BELONGS TO THE S8E FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; Z99259; CAB16376.1; -
DR      InterPro; IPR001047; Ribosomal_S8E.
DR      Pfam; PF01201; Ribosomal_S8e; 1.
DR      ProDom; PD005658; Ribosomal_S8E; 1.
DR      TIGRFAMS; TIGR00307; S8e; 1.
DR      PROSITE; PS01193; RIBOSOMAL_S8E; 1.
KW      Ribosomal protein; Multigene family.
SQ      SEQUENCE 200 AA; 22593 MW; F03C815C75E35E3B CRC64;

Query Match      100.0%; Score 21; DB 1; Length 200;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPF 4
        ||||
Db      106 AAPF 109

RESULT 37
RS8B_SCHPO
ID  RS8B_SCHPO      STANDARD;      PRT;      200 AA.
AC  Q9P7B2;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  40S ribosomal protein S8-B.
GN  RPS8B OR SPAC521.05.
OS  Schizosaccharomyces pombe (Fission yeast).
OC  Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC  Schizosaccharomycetales; Schizosaccharomycetaceae;
OC  Schizosaccharomyces.
OX  NCBI_TaxID=4896;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=972;
RX  MEDLINE=21848401; Pubmed=11859360;
RA  Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA  Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA  Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA  Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA  Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA  Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA  James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA  Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA  Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA  Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA  Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA  Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA  Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA  Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA  Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA  Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA  Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA  Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA  Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA  Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA  Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA  Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA  Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA  Shpakovski G.V., Ussey D., Barrell B.G., Nurse P.;
RT  "The genome sequence of Schizosaccharomyces pombe.";
RL  Nature 415:871-880(2002).
CC  -1- MISCELLANEOUS: THERE ARE TWO GENES FOR S8 IN S.POMBE.
CC  -1- SIMILARITY: BELONGS TO THE S8E FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
```



```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL163481; CAB86469.1; -.
DR InterPro; IPR001047; Ribosomal_S8E.
DR Pfam; PF01201; Ribosomal_S8E; 1.
DR ProDom; PD005658; Ribosomal_S8E; 1.
DR TIGRExams; TIGR00307; S8e; 1.
DR PROSITE; PS01193; RIBOSOMAL_S8E; 1.
DR Ribosomal protein; Multigene family.
SQ SEQUENCE 200 AA; 22564 MW; 282A4D93800353B6 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 200;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
    ||||
Db 106 AAPF 109

RESULT 38
YD99_HAEIN STANDARD; PRT; 202 AA.
ID YD99_HAEIN
AC P4175;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein H11399.
GN H11399.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OX Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."
RL Science 269:496-512(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32820; AAC23051.1; -.
DR TIGR; H11399; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 202 AA; 23238 MW; D55017DAB9D191C4 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 202;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
    ||||
DR
```

```
Db 16 AAPF 19

RESULT 39
YWRF_BACSU STANDARD; PRT; 205 AA.
ID YWRF_BACSU
AC 005220;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ywrf.
GN YWRF.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98015417; PubMed=9353933;
RA Kunz F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriis R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Chai S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambuit R., Wedler E., Wedler H., Weltzenegger T.,
RA Yoshida K., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -i- COFACTOR: FMN (By similarity).
CC -i- SIMILARITY: BELONGS TO THE FLAVOREDOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z93767; CAB07792.1; -.
DR EMBL; Z99122; CAB15625.1; -.
CC
```


DR HSSP; 026255; 1EJE.
 DR Subtilist; BG12525; ywrf.
 KW Hypothetical protein; Flavoprotein; FMN; Complete proteome.
 SQ SEQUENCE 205 AA; 22584 MW; 28A3A8F5CC4529F CRC64;

Query Match 100.0%; Score 21; DB 1; Length 205;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
 1111
 Db 41 AAPF 44

RESULT 40
 YM80_YEAST

ID YM80_YEAST STANDARD; PRT; 206 AA.
 AC 004019;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 23.2 kDa protein in ZRC1-FAA4 intergenic region.
 GN YMR245W OR YMR408.07.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;

[1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288c / AB972;
 RA Gentles S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Z48756; CAA88655.1; -
 DR SGD; S0004859; YMR245W.
 KW Hypothetical protein.
 SQ SEQUENCE 206 AA; 23194 MW; 8794E591DAB49C6E CRC64;

Query Match 100.0%; Score 21; DB 1; Length 206;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
 1111
 Db 138 AAPF 141

Search completed: December 6, 2002, 13:28:33
 Job time : 6.55556 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:25:09 ; Search time 19.5556 Seconds
(without alignments)
42.146 Million cell updates/sec

Title: US-10-033-526-1
Perfect score: 21
Sequence: 1 AAPF 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|---------------------|
| 1 | 21 | 100.0 | 41 | 10 | Q9S9F1 | Q9S9F1 brassica na |
| 2 | 21 | 100.0 | 45 | 2 | P70957 | P70957 bacillus st |
| 3 | 21 | 100.0 | 47 | 6 | O9TREG | O9TREG sus scrofa |
| 4 | 21 | 100.0 | 47 | 6 | O9TQX9 | O9TQX9 bos taurus |
| 5 | 21 | 100.0 | 50 | 16 | O8YCK3 | O8YCK3 bruceella me |
| 6 | 21 | 100.0 | 51 | 4 | O9Y5N3 | O9Y5N3 homo sapien |
| 7 | 21 | 100.0 | 57 | 10 | O9M7N1 | O9M7N1 triticum ae |
| 8 | 21 | 100.0 | 59 | 16 | O9ADG2 | O9ADG2 streptomyc |
| 9 | 21 | 100.0 | 69 | 10 | O8S4B7 | O8S4B7 amblystegiu |
| 10 | 21 | 100.0 | 70 | 12 | O9DWS4 | O9DWS4 hepatitis c |
| 11 | 21 | 100.0 | 72 | 7 | O9MWJ2 | O9MWJ2 certhidea o |
| 12 | 21 | 100.0 | 73 | 5 | O26512 | O26512 schistosoma |
| 13 | 21 | 100.0 | 75 | 16 | O8UEW5 | O8UEW5 agrobacteri |
| 14 | 21 | 100.0 | 76 | 13 | O9DFN9 | O9DFN9 gilllichthys |
| 15 | 21 | 100.0 | 77 | 16 | O98167 | O98167 rhizobium 1 |
| 16 | 21 | 100.0 | 78 | 10 | O8S493 | O8S493 amblystegiu |

| | | | | | | |
|----|----|-------|----|----|--------|--------------------|
| 17 | 21 | 100.0 | 79 | 10 | O8S4B6 | O8S4B6 amblystegiu |
| 18 | 21 | 100.0 | 79 | 10 | O8S4B4 | O8S4B4 amblystegiu |
| 19 | 21 | 100.0 | 79 | 10 | O8S4B2 | O8S4B2 amblystegiu |
| 20 | 21 | 100.0 | 79 | 10 | O8S4A7 | O8S4A7 amblystegiu |
| 21 | 21 | 100.0 | 79 | 10 | O8S499 | O8S499 amblystegiu |
| 22 | 21 | 100.0 | 79 | 17 | O26308 | O26308 methanobact |
| 23 | 21 | 100.0 | 80 | 16 | O971W6 | O971W6 clostridium |
| 24 | 21 | 100.0 | 81 | 2 | O51347 | O51347 pseudomonas |
| 25 | 21 | 100.0 | 81 | 7 | P79449 | P79449 bos taurus |
| 26 | 21 | 100.0 | 81 | 7 | P79450 | P79450 bos taurus |
| 27 | 21 | 100.0 | 83 | 12 | O9DKP9 | O9DKP9 hepatitis b |
| 28 | 21 | 100.0 | 86 | 16 | O8X3A0 | O8X3A0 escherichia |
| 29 | 21 | 100.0 | 87 | 10 | O9FSP9 | O9FSP9 oryza sativ |
| 30 | 21 | 100.0 | 88 | 16 | O8UH09 | O8UH09 agrobacteri |
| 31 | 21 | 100.0 | 90 | 2 | O9Z3D9 | O9Z3D9 escherichia |
| 32 | 21 | 100.0 | 90 | 5 | O95NY3 | O95NY3 drosophila |
| 33 | 21 | 100.0 | 90 | 5 | O95NY2 | O95NY2 drosophila |
| 34 | 21 | 100.0 | 90 | 5 | O95NY1 | O95NY1 drosophila |
| 35 | 21 | 100.0 | 90 | 5 | O97184 | O97184 drosophila |
| 36 | 21 | 100.0 | 90 | 5 | O97185 | O97185 drosophila |
| 37 | 21 | 100.0 | 90 | 5 | O97456 | O97456 drosophila |
| 38 | 21 | 100.0 | 90 | 5 | O97460 | O97460 drosophila |
| 39 | 21 | 100.0 | 92 | 2 | O9ANK5 | O9ANK5 bradyrhizob |
| 40 | 21 | 100.0 | 92 | 7 | O95470 | O95470 bos indicus |
| 41 | 21 | 100.0 | 94 | 16 | O8U4Y1 | O8U4Y1 agrobacteri |
| 42 | 21 | 100.0 | 96 | 10 | O8S495 | O8S495 amblystegiu |
| 43 | 21 | 100.0 | 97 | 10 | O8S4B1 | O8S4B1 amblystegiu |
| 44 | 21 | 100.0 | 97 | 11 | O9CVJ8 | O9CVJ8 mus musculu |
| 45 | 21 | 100.0 | 97 | 16 | O92EA6 | O92EA6 listeria in |

ALIGNMENTS

RESULT 1

Q9S9F1 ID Q9S9F1 PRELIMINARY; PRT; 41 AA.
AC Q9S9F1; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE Napin short chain S4-CALMODULIN antagonist/calcium-dependent protein
DE kinase substrate.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE.
RX MEDLINE=96283790; PubMed=8679670;
RA Neumann G.M., Condron R., Thomas I., Polya G.M.;
RT "Purification and sequencing of multiple forms of Brassica napus seed
RT napin small chains that are calmodulin antagonists and substrates for
RT plant calcium-dependent protein kinase.";
RL Biochim. Biophys. Acta 1295:23-33(1996).
SQ SEQUENCE 41 AA; 4652 MW; 8BC6738503380553 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 41;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 2 AAPF 5

RESULT 2
P70957 ID P70957 PRELIMINARY; PRT; 45 AA.
AC P70957; 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)

DT 01-NOV-1998 (TREMBLrel. 08, last annotation update)
DE Alpha-amylase precursor (Fragment).
GN AMYS.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX Geobacillus.
RN NCBI_TaxID=1422;
RP [1]
RX SEQUENCE FROM N.A.
MEDLINE=91092499; PubMed=2265757;
RA Jorgensen P.L., Hansen C.K., Poulsen G.B., Diderichsen B.;
RT "In vivo genetic engineering: homologous recombination as a tool for
plasmid construction."
RL Gene 96:37-41(1990).
DR EMBL; M62638; AAA22242.1; -.
KW Signal.
FT SIGNAL. 1 34 POTENTIAL.
FT CHAIN 35 >45 POTENTIAL.
FT NON_TER 45 45
SQ SEQUENCE 45 AA; 5119 MW; 192F623EA13E6DA6 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
| | | |
Db 35 AAPF 38

RESULT 3

O9TRE9 PRELIMINARY; PRT; 47 AA.
AC O9TRE9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Cytochrome c oxidase subunit VIIC (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RX MEDLINE=93384597; PubMed=8396926;
RA Sillard R., Jornvall H., Mutt V.;
RT "Characterization of porcine intestinal cytochrome c oxidase subunit
VIIC, purified by affinity chromatography."
RL Biochem. Biophys. Res. Commun. 195:746-750(1993).
DR InterPro; IPR004202; COX7C.
DR Pfam; PF02935; COX7C; 1.
SQ SEQUENCE 47 AA; 5457 MW; E11E89CFBC18B5D7 CRC64;

Query Match 100.0%; Score 21; DB 6; Length 47;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
| | | |
Db 34 AAPF 37

RESULT 4

O9TOX9 PRELIMINARY; PRT; 47 AA.
ID O9TOX9;
AC O9TOX9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Deoxyribonuclease F (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=94183469; PubMed=8136077;
RA Chang Y.M., Lin S., Liao T.H.;
RT "Bovine pancreatic deoxyribonuclease F: isoelectric focusing, peptide
mapping and primary structure."
RL Biotechnol. Appl. Biochem. 19:129-140(1994).
DR HSSP; P00639; 3DNI.
DR InterPro; IPR01582; DNase_I.
DR PRINTS; PR00130; DNASEI.
DR ProDom; PD005408; DNase_I; 1.
SQ SEQUENCE 47 AA; 4937 MW; D757A9F441095799 CRC64;

Query Match 100.0%; Score 21; DB 6; Length 47;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
| | | |
Db 17 AAPF 20

RESULT 5

O8YCK3 PRELIMINARY; PRT; 50 AA.
ID O8YCK3
AC O8YCK3;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Hypothetical protein BMEI10525.
GN BMEI10525.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
RX MEDLINE=20020109; PubMed=11756688;
RA Delvecchio V.G., Kapatral V., Redkar R.J., Patra G., Mufer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyripides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
Brucella melitensis."
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009688; AAL53767.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 50 AA; 5208 MW; 8DE59A82243EFC17 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
| | | |
Db 44 AAPF 47

RESULT 6

O9Y5N3 PRELIMINARY; PRT; 51 AA.
ID O9Y5N3
AC O9Y5N3;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Lecithin-cholesterol acyltransferase (EC 2.3.1.43) (Fragment).
GN LCAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ou J.F., Saku K., Liao Y.L., Jimi S., Jin W.J., Arakawa K.;
RT "A novel missense variant Thr13met in the lecithin: cholesterol
RT acyltransferase (LCAT) gene associated with fish eye disease.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF140503; AAD28484.1; -.
KW Acyltransferase; Transferase.
FT VARIANT 37 37 M -> T.
FT NON_TER 51 51
SQ SEQUENCE 51 AA; 5666 MW; EB511439A5129965 CRC64;

Query Match 100.0%; Score 21; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   ||||
Db 22 AAPF 25

RESULT 7
O9M7N1
ID O9M7N1 PRELIMINARY; PRT; 57 AA.
AC O9M7N1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MFP1 attachment factor 1 (Fragment).
GN MFP1.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99418887; Pubmed=10488241;
RA Gindullis F., Peffer N.J., Meier I.;
RT "MFP1, a novel plant protein interacting with matrix attachment region
RT binding protein MFP1, is located at the nuclear envelope.";
RL Plant Cell 11:1755-1768(1999).
DR EMBL; AF118118; AAF63662.1; -.
FT NON_TER 57 57
SQ SEQUENCE 57 AA; 6085 MW; E35F7AC99E250F78 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 57;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   ||||
Db 17 AAPF 20

RESULT 8
O9ADG2
ID O9ADG2 PRELIMINARY; PRT; 59 AA.
AC O9ADG2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein SCO6101.
GN SCO6101 OR SCBAC1A6.25C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
```

```
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL589708; CAC33946.1; -.
KW Hypothetical protein.
SQ SEQUENCE 59 AA; 6387 MW; 885755E65DB6C9EF CRC64;

Query Match 100.0%; Score 21; DB 16; Length 59;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   ||||
Db 5 AAPF 8

RESULT 9
O8S4B7
ID O8S4B7 PRELIMINARY; PRT; 69 AA.
AC O8S4B7;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Adenosine kinase (Fragment).
OS Amblystegium humile.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Bryidae; Hypnales; Amblystegiaceae; Amblystegium.
OX NCBI_TaxID=111433;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AH2C;
RA Vanderpoorten A., Shaw A.J., Cox C.J.;
RT "Reconciled gene trees, molecular evolution, and species relationships
RT in Amblystegium.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF465020; AAL96441.1; -.
KW kinase.
FT NON_TER 69 69
FT NON_TER 1 1
SQ SEQUENCE 69 AA; 7793 MW; 81C1C22B4EAC30AC CRC64;

Query Match 100.0%; Score 21; DB 10; Length 69;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   ||||
Db 40 AAPF 43

RESULT 10
O9DWS4
ID O9DWS4 PRELIMINARY; PRT; 70 AA.
AC O9DWS4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E1 (GP32) (GP35)]
DE (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
```

RP SEQUENCE FROM N.A.
RC STRAIN-VE3216;
RX MEDLINE=20394384; PubMed=10935993;
RA Argentin C., Dettori S., Villano U., Guadagnino V., Infantolino D.,
RA Dentico P., Coppola R.C., Rapicetta M.;
RT "Molecular characterisation of HCV genotype 4 isolates circulating in
Italy.";
RL J. Med. Virol. 62:84-90(2000).
CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR EMBL; AJ250216; CAC16883.1; -.
DR InterPro; IPR002519; HCV_env.
DR Pfam; PF01539; HCV_env; 1.
KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 70 70
SQ SEQUENCE 70 AA; 7698 MW; BD360B781C3BF80F CRC64;

Query Match 100.0%; Score 21; DB 12; Length 70;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 54 AAPF 57

RESULT 11
Q9MWJ2 PRELIMINARY; PRT; 72 AA.
ID Q9MWJ2;
AC Q9MWJ2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE MHC class IIB antigen (Fragment).
DE Certhidea olivacea.
OS Certhidea olivacea.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;
OC Fringillidae; Emberizinae; Certhidea.
OX NCBI_TaxID=48880;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D10ARF;
RA Sato A., Figueroa F., Mayer W.E., Grant P.R., Grant R., Klein J.;
RT "MHC class II genes of Darwin's finches: divergence by point mutations
and reciprocal recombination.";
RL (in) Kasahara M. (eds.);
RL Major histocompatibility complex:
RL evolution, structure, and function, pp.518-541, Springer-Verlag,
RL Tokyo (2000).
RL EMBL; AF164163; AAF36225.1; -.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.
DR Prodom; PD000328; MHC_II_beta; 1.
FT NON_TER 1 1
FT NON_TER 72 72
SQ SEQUENCE 72 AA; 8760 MW; F62FEF27A7B4A672B CRC64;

Query Match 100.0%; Score 21; DB 7; Length 72;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 63 AAPF 66

RESULT 12
Q26512 PRELIMINARY; PRT; 73 AA.
ID Q26512;
AC Q26512;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Triosephosphate isomerase (Fragment).
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidae;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHINESE AND PHILIPPINE STRAINS;
RX MEDLINE=97101875; PubMed=8946393;
RA Hooker C.W., Brindley P.J.;
RT "Cloning and characterisation of strain-specific transcripts encoding
triosephosphate isomerase, a candidate vaccine antigen from
Schistosoma japonicum.";
RL MOL. Biochem. Parasitol. 82:265-269(1996).
DR EMBL; U57557; AAC47394.1; -.
DR HSSP; P00940; TIM.
DR InterPro; IPR000652; Triophos_ismrse.
DR Pfam; PF00121; TIM; 1.
DR Prodom; PD001005; Triophos_ismrse; 1.
KW Isomerase.
FT NON_TER 1 1
FT NON_TER 73 73
SQ SEQUENCE 73 AA; 7982 MW; 6B10DB35855F5A33 CRC64;

Query Match 100.0%; Score 21; DB 5; Length 73;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 35 AAPF 38

RESULT 13
Q8UEW5 PRELIMINARY; PRT; 75 AA.
ID Q8UEW5;
AC Q8UEW5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein Atu1639.
GN ATU1639.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutlyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58.";
RL Science 294:2317-2323(2001).
DR EMBL; AE009121; AAL42640.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 75 AA; 8397 MW; B3B9D20CE8EE82DD CRC64;

Query Match 100.0%; Score 21; DB 16; Length 75;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4

Db 50 AAPF 53

RESULT 14

09DFN9 PRELIMINARY; PRT; 76 AA.
AC Q9DFN9;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Formyltetrahydrofolate dehydrogenase (Fragment).
OS Gillichthys mirabilis (Long-jawed mudsucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Gobioidi;
OC Gobiidae; Gillichthys.
OX NCBI_TaxID=8222;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=21117151; PubMed=11172064;
RA Gracey A.Y., Troil J.V., Somero G.N.;
RT "Hypoxia-induced gene expression profiling in the euryoxic fish
Gillichthys mirabilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:1993-1998(2001).
DR EMBL; AF266196; AAG13316.1; -.
DR HSSP; P05091; 1CW3.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldedh; 1.
FT NON_TER 1
SQ SEQUENCE 76 AA; 8182 MW; 787505071A1ACFA3 CRC64;

Query Match

Best Local Similarity 100.0%; Score 21; DB 13; Length 76;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4

Db 43 AAPF 46

RESULT 15

098L67 PRELIMINARY; PRT; 77 AA.
AC Q98L67;
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein msr1158.
GN MSRI158.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP002996; BAB48596.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 77 AA; 8189 MW; 8A2BF074EAC16806 CRC64;

Query Match

Best Local Similarity 100.0%; Score 21; DB 16; Length 77;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 54 AAPF 57

RESULT 16

08S493 PRELIMINARY; PRT; 78 AA.
AC Q8S493;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Adenosine kinase (Fragment).
OS Amblystegium fluviatile.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Bryidae; Hypnales; Amblystegiaceae; Amblystegium.
OX NCBI_TaxID=140014;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2A;
RA Vanderpoorten A., Shaw A.J., Cox C.J.;
RT "Reconciled gene trees, molecular evolution, and species
relationships in Amblystegium (Bryopsida).";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF465154; AAL96478.1; -.
KW kinase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 78 AA; 8863 MW; 10D283D076391CBC CRC64;

Query Match

Best Local Similarity 100.0%; Score 21; DB 10; Length 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4

Db 40 AAPF 43

RESULT 17

08S4B6 PRELIMINARY; PRT; 79 AA.
AC Q8S4B6;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Adenosine kinase (Fragment).
OS Amblystegium humile.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Bryidae; Hypnales; Amblystegiaceae; Amblystegium.
OX NCBI_TaxID=111433;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AH2E;
RA Vanderpoorten A., Shaw A.J., Cox C.J.;
RT "Reconciled gene trees, molecular evolution, and species relationships
in Amblystegium.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF465022; AAL96443.1; -.
KW kinase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 79 AA; 9026 MW; 9520D283C42A091C CRC64;

Query Match

Best Local Similarity 100.0%; Score 21; DB 10; Length 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4

Db 40 AAPF 43

```
RESULT 18
O8S4B4
ID O8S4B4 PRELIMINARY; PRT; 79 AA.
AC O8S4B4;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Adenosine kinase (Fragment).
OS Amblystegium serpens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Bryidae; Hypnales; Amblystegiaceae; Amblystegium.
OX NCBI_TaxID=99400;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AHPRIMD;
RA Vanderpoorten A., Shaw A.J., Cox C.J.;
RT "Reconciled gene trees, molecular evolution, and species relationships
in Amblystegium."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF465024; AAL96445.1; -.
KW kinase.
FT NON_TER 1 1
FT NON_TER 79 79
SQ SEQUENCE 79 AA; 9010 MW; 9520D283D07A091C CRC64;
```

```
Query Match 100.0%; Score 21; DB 10; Length 79;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPF 4
    |||
Db 40 AAPF 43
```

```
RESULT 19
O8S4B2
ID O8S4B2 PRELIMINARY; PRT; 79 AA.
AC O8S4B2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Adenosine kinase (Fragment).
OS Amblystegium humile.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Bryidae; Hypnales; Amblystegiaceae; Amblystegium.
OX NCBI_TaxID=111433;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AHSECB;
RA Vanderpoorten A., Shaw A.J., Cox C.J.;
RT "Reconciled gene trees, molecular evolution, and species relationships
in Amblystegium."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF465026; AAL96447.1; -.
KW kinase.
FT NON_TER 1 1
FT NON_TER 79 79
SQ SEQUENCE 79 AA; 9010 MW; 9520D283D07A091C CRC64;
```

```
Query Match 100.0%; Score 21; DB 10; Length 79;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPF 4
    |||
Db 40 AAPF 43
```

```
RESULT 20
O8S4A7
ID O8S4A7 PRELIMINARY; PRT; 79 AA.
AC O8S4A7;
```

```
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Adenosine kinase (Fragment).
OS Amblystegium varium.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Bryidae; Hypnales; Amblystegiaceae; Amblystegium.
OX NCBI_TaxID=111436;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AVC;
RA Vanderpoorten A., Shaw A.J., Cox C.J.;
RT "Reconciled gene trees, molecular evolution, and species relationships
in Amblystegium."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF465032; AAL96453.1; -.
KW kinase.
FT NON_TER 1 1
FT NON_TER 79 79
SQ SEQUENCE 79 AA; 8950 MW; 9530C382C06B081C CRC64;
```

```
Query Match 100.0%; Score 21; DB 10; Length 79;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPF 4
    |||
Db 40 AAPF 43
```

```
RESULT 21
O8S499
ID O8S499 PRELIMINARY; PRT; 79 AA.
AC O8S499;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Adenosine kinase (Fragment).
OS Amblystegium varium.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Bryidae; Hypnales; Amblystegiaceae; Amblystegium.
OX NCBI_TaxID=111436;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AVSECB;
RA Vanderpoorten A., Shaw A.J., Cox C.J.;
RT "Reconciled gene trees, molecular evolution, and species
relationships in Amblystegium (Bryopsida).";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF465146; AAL96470.1; -.
KW kinase.
FT NON_TER 1 1
FT NON_TER 79 79
SQ SEQUENCE 79 AA; 9010 MW; 9520D283D07A091C CRC64;
```

```
Query Match 100.0%; Score 21; DB 10; Length 79;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AAPF 4
    |||
Db 40 AAPF 43
```

```
RESULT 22
O26308
ID O26308 PRELIMINARY; PRT; 79 AA.
AC O26308;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein MTH206.
GN MTH206.
```

OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=93711463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzaden R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL; AE000807; AAB84712.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 79 AA; 9126 MW; F55F5EAB36C31635 CRC64;

Query Match 100.0%; Score 21; DB 17; Length 79;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 36 AAPF 39

RESULT 23

O97LM6 PRELIMINARY; PRT; 80 AA.
AC O97LM6;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Predicted metal-binding protein.
GN CAC0438.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007558; AAK78418.1; -;
KW Complete proteome.
SQ SEQUENCE 80 AA; 9080 MW; 9C900BD213AC22BB CRC64;

Query Match 100.0%; Score 21; DB 16; Length 80;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 63 AAPF 66

RESULT 24

O51347 PRELIMINARY; PRT; 81 AA.
AC O51347;
DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 8.5 kDa protein.
OS Pseudomonas aeruginosa.
OG Plasmid RI033.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-IS6100DELTA;
RX MEDLINE=89364699; PubMed=2549372;
RA Wohlleben W., Arnold W., Bissonnette L., Pelletier A., Tanguay A.,
RA Roy P.H., Gamboa G.C., Barry G.F., Aubert E., Davies J., Kagan S.A.;
RT "On the evolution of Tn21-like multiresistance transposons: sequence
analysis of the gene (aacC1) for gentamicin acetyltransferase-3-
I(AAC(3)-I), another member of the Tn21-based expression cassette.";
RT I(AAC(3)-I), another member of the Tn21-based expression cassette.";
RL Mol. Gen. Genet. 217:202-208(1989).
DR EMBL; U12338; AAB60001.1; -;
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 81 AA; 8524 MW; 61C21CF47F15EC50 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 49 AAPF 52

RESULT 25

P79449 PRELIMINARY; PRT; 81 AA.
AC P79449;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Major histocompatibility complex class II (Fragment).
GN DQB1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92165349; PubMed=1537611;
RA Sigurdardottir S., Borsch C., Gustafsson K., Andersson L.;
RT "Gene duplications and sequence polymorphism of bovine class II DQB
genes.";
RT Immunogenetics 35:205-213(1992).
RN [2]
RP SEQUENCE FROM N.A.

RA Mikko S.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U77793; AAB39395.1; -;
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1
FT NON_TER 81
FT NON_TER 81
SQ SEQUENCE 81 AA; 9727 MW; D8F1F1663F4187DD CRC64;

Query Match 100.0%; Score 21; DB 7; Length 81;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
Db 78 AAPF 81

```
RESULT 26
P79450
ID P79450 PRELIMINARY; PRT; 81 AA.
AC P79450;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Major histocompatibility complex class II (Fragment).
GN DQB1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92165349; PubMed=1537611;
RA Sigurdardottir S., Borsch C., Gustafsson K., Andersson L.;
RT "Gene duplications and sequence polymorphism of bovine class II DQB
RT genes."
RL Immunogenetics 35:205-213(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Mikko S.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U77794; AAB39396.1; -.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1
FT NON_TER 81
SQ SEQUENCE 81 AA; 9808 MW; 7BDF9A930A6FEA1B6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 7; Length 81;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 78 AAPF 81

RESULT 27
Q9DKP9
ID Q9DKP9 PRELIMINARY; PRT; 83 AA.
AC Q9DKP9;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE Mutant polymerase (Fragment).
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Dong J., Cheng J., Wang Q., Shi S., Zhang J., Xia X., Si C.;
RT "HBV quasiespecies: RT region and S gene mutant."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF329860; AAG48741.1; -.
DR InterPro; IPR001462; DNAPol_viral_C.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR ProDom; PD000814; DNAPol_viral_C; 1.
FT NON_TER 1
FT NON_TER 83
SQ SEQUENCE 83 AA; 9409 MW; 48C125F9A0657A13 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 12; Length 83;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
```

```
Db 48 AAPF 51

RESULT 28
O8X3A0
ID O8X3A0 PRELIMINARY; PRT; 86 AA.
AC O8X3A0;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein ECS2213.
GN ECS2213.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
DR EMBL; AP002557; BAB35636.1; -.
KW Hypothetical protein.
SQ SEQUENCE 86 AA; 9659 MW; B034A45AFCCF137F CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 16; Length 86;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 26 AAPF 29

RESULT 29
Q9FSP9
ID Q9FSP9 PRELIMINARY; PRT; 87 AA.
AC Q9FSP9;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Hypothetical 9.8 kDa protein.
GN H0711G06.1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Mu J., Zhou B., Chen Z.H., Li Y., Zhu J.J.,
RA Tang Y.S., Zhao Q., Liu Y.L., Huang Y.C., Yu Z., Fan D.L., Chen L.,
RA Weng Q.J., Zhang L., Lu Y.Q., Yu S.L., Zhu J., Liu X.H., Hu X.,
RA Lei H.Y., Zhang Y.J., Wang R., Li C., Lu Y., Chen X.C., Zhang Y.,
RA Hu H., Jia P.X., Li T., Qian Y.M., Ying K., Hong G.F.;
RT "Oryza sativa indica (Guangluai4) genomic DNA, chromosome 4, BAC
RT clone: H0711G06 (+H0113C06).";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL442115; CAC09495.1; -.
KW Hypothetical protein.
SQ SEQUENCE 87 AA; 9791 MW; BE347F35755D6A06 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 87;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 AAPF 4
1111
Db 43 AAPF 46

RESULT 30
O8UH09
ID O8UH09 PRELIMINARY; PRT; 88 AA.
AC O8UH09;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE Flagellar biosynthetic protein.
GN FLIQ OR ATU0580 OR AGR_C_1023.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P., Woo L., Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Wu L., Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr., Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C., Kuttyavin T., Levy R., Li M.-J., McClelland E., Palmeri A., Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D., Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M., Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M., Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V., Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L., Houmlet K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F., Woliam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B., Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G., Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE009026; AAL41597.1; -.
DR EMBL; AE007992; AAK86391.1; -.
KW Complete proteome.
SQ SEQUENCE 88 AA; 9234 MW; AF1C2039D903160F CRC64;

Query Match 100.0%; Score 21; DB 16; Length 88;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
Db 65 AAPF 68

RESULT 31
O9Z3D9
ID O9Z3D9 PRELIMINARY; PRT; 90 AA.
AC O9Z3D9;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE Hypothetical 24.1 kDa protein in LEF4-P33 intergenic region.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A., Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K., Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N., Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
DR EMBL; D90752; BAA36020.1; -.
DR EMBL; D90751; BAA36007.1; -.
SQ SEQUENCE 90 AA; 9470 MW; A3F8A7CA3F662ABA CRC64;

Query Match 100.0%; Score 21; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
Db 33 AAPF 36

RESULT 32
O95NY3
ID O95NY3 PRELIMINARY; PRT; 90 AA.
AC O95NY3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE Acp26Ab protein.
GN ACP26AB OR CG9024.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LA27, MO13A, AND MO34A;
RX MEDLINE=99016087; PubMed=9799260;
RA Aguade M.;
RT "Different forces drive the evolution of the Acp26Aa and Acp26Ab RT accessory gland genes in the Drosophila melanogaster species RT complex.";
RL Genetics 150:1079-1089(1998).
DR EMBL; AJ231366; CAB37227.1; -.
DR EMBL; AJ231350; CAB37195.1; -.
DR EMBL; AJ231352; CAB37199.1; -.
DR FlyBase; FBgn0002856; Acp26Ab.
SQ SEQUENCE 90 AA; 10148 MW; 1FDFC4DA51F9BDF CRC64;

Query Match 100.0%; Score 21; DB 5; Length 90;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
1111
Db 21 AAPF 24

RESULT 33
O95NY2
ID O95NY2 PRELIMINARY; PRT; 90 AA.
AC O95NY2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE Acp26Ab protein.
GN ACP26AB OR CG9024.
OS Drosophila melanogaster (Fruit fly).


```
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LA58, LA37, LA54, MA11, MA20, MA35, AND MA53;
RX MEDLINE=99016087; PubMed=9799260;
RA Aguade M.;
RT "Different forces drive the evolution of the Acp26Aa and Acp26Ab
RT accessory gland genes in the Drosophila melanogaster species
RT complex.";
RL Genetics 150:1079-1089(1998).
DR EMBL; AJ231374; CAB37624.1; -.
DR EMBL; AJ231370; CAB37235.1; -.
DR EMBL; AJ231371; CAB37237.1; -.
DR EMBL; AJ231373; CAB37241.1; -.
DR EMBL; AJ231386; CAB37265.1; -.
DR EMBL; AJ231388; CAB37269.1; -.
DR EMBL; AJ231393; CAB37279.1; -.
DR EMBL; AJ231397; CAB37287.1; -.
DR FLYBase; FBgn0002856; Acp26Ab.
SQ SEQUENCE 90 AA; 10171 MW; 1FDFC4E8D8F9BDFF CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 90;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 21 AAPF 24

RESULT 34
ID Q95NY1 PRELIMINARY; PRT; 90 AA.
AC Q95NY1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE ACP26Ab protein.
GN ACP26AB OR CG9024.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VARIOUS STRAINS;
RX MEDLINE=99016087; PubMed=9799260;
RA Aguade M.;
RT "Different forces drive the evolution of the Acp26Aa and Acp26Ab
RT accessory gland genes in the Drosophila melanogaster species
RT complex.";
RL Genetics 150:1079-1089(1998).
DR EMBL; AJ231395; CAB37283.1; -.
DR EMBL; AJ231351; CAB37197.1; -.
DR EMBL; AJ231355; CAB37205.1; -.
DR EMBL; AJ231357; CAB37209.1; -.
DR EMBL; AJ231360; CAB37215.1; -.
DR EMBL; AJ231361; CAB37217.1; -.
DR EMBL; AJ231362; CAB37219.1; -.
DR EMBL; AJ231364; CAB37223.1; -.
DR EMBL; AJ231376; CAB37245.1; -.
DR EMBL; AJ231377; CAB37247.1; -.
DR EMBL; AJ231380; CAB37253.1; -.
DR EMBL; AJ231382; CAB37257.1; -.
DR EMBL; AJ231384; CAB37261.1; -.
DR EMBL; AJ231387; CAB37267.1; -.
DR EMBL; AJ231390; CAB37273.1; -.
DR FLYBase; FBgn0002856; Acp26Ab.
SQ SEQUENCE 90 AA; 10157 MW; 1FD4DC4BD8F9BDFF CRC64;
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 90;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 21 AAPF 24

RESULT 35
ID Q97184 PRELIMINARY; PRT; 90 AA.
AC Q97184;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE ACP26AB protein.
GN ACP26AB OR CG9024.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LA28;
RX MEDLINE=99016087; PubMed=9799260;
RA Aguade M.;
RT "Different forces drive the evolution of the Acp26Aa and Acp26Ab
RT accessory gland genes in the Drosophila melanogaster species
RT complex.";
RL Genetics 150:1079-1089(1998).
DR EMBL; AJ231367; CAB37229.1; -.
DR FLYBase; FBgn0002856; Acp26Ab.
SQ SEQUENCE 90 AA; 10171 MW; 1FD4C0B9D8F9BDFF CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 90;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
Db 21 AAPF 24

RESULT 36
ID Q97185 PRELIMINARY; PRT; 90 AA.
AC Q97185;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE ACP26AB protein.
GN ACP26AB OR CG9024.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA60;
RX MEDLINE=99016087; PubMed=9799260;
RA Aguade M.;
RT "Different forces drive the evolution of the Acp26Aa and Acp26Ab
RT accessory gland genes in the Drosophila melanogaster species
RT complex.";
RL Genetics 150:1079-1089(1998).
DR EMBL; AJ231400; CAB37293.1; -.
DR FLYBase; FBgn0002856; Acp26Ab.
SQ SEQUENCE 90 AA; 10139 MW; 1FD91C4BD8F9BDFF CRC64;
```

```
Query Match          100.0%; Score 21; DB 5; Length 90;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   |||
Db 21 AAPF 24

RESULT 37
O97456 PRELIMINARY; PRT; 90 AA.
AC O97456;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE ACP26AB protein.
GN ACP26AB OR CG9024.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA74, LA31, LA106, LA118, MA6, MA31, MA37, MA56, AND MA57;
RX MEDLINE=99016087; Pubmed=9799260;
RA Aguade M.;
RT "Different forces drive the evolution of the Acp26Aa and Acp26Ab
RT accessory gland genes in the Drosophila melanogaster species
RT complex.";
RL Genetics 150:1079-1089(1998).
DR EMBL; AJ231401; CAB37629.1; -
DR EMBL; AJ231368; CAB37231.1; -
DR EMBL; AJ231378; CAB37249.1; -
DR EMBL; AJ231381; CAB37255.1; -
DR EMBL; AJ231385; CAB37263.1; -
DR EMBL; AJ231392; CAB37277.1; -
DR EMBL; AJ231394; CAB37281.1; -
DR EMBL; AJ231398; CAB37289.1; -
DR EMBL; AJ231399; CAB37291.1; -
DR FLYBase; FBgn0002856; Acp26Ab.
SQ SEQUENCE 90 AA; 10215 MW; 191FDC4BD8F9BDFF CRC64;

Query Match          100.0%; Score 21; DB 5; Length 90;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   |||
Db 21 AAPF 24

RESULT 38
O97460 PRELIMINARY; PRT; 90 AA.
AC O97460;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE ACP26AB protein.
GN ACP26AB OR CG9024.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA50, LA14, LA32, LA60, LA125, MA21, AND MA24;
RX MEDLINE=99016087; Pubmed=9799260;
RA Aguade M.;
RT "Different forces drive the evolution of the Acp26Aa and Acp26Ab
RT accessory gland genes in the Drosophila melanogaster species
RT complex.";
RL Genetics 150:1079-1089(1998).
DR EMBL; AJ231396; CAB37285.1; -
DR EMBL; AJ231363; CAB37221.1; -
DR EMBL; AJ231369; CAB37233.1; -
DR EMBL; AJ231375; CAB37243.1; -
DR EMBL; AJ231383; CAB37259.1; -
DR EMBL; AJ231389; CAB37271.1; -
DR EMBL; AJ231391; CAB37275.1; -
DR FLYBase; FBgn0002856; Acp26Ab.
SQ SEQUENCE 90 AA; 10229 MW; 191FC4E8D8F9BDFF CRC64;

Query Match          100.0%; Score 21; DB 5; Length 90;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   |||
Db 21 AAPF 24

RESULT 39
O9ANK5 PRELIMINARY; PRT; 92 AA.
ID O9ANK5
AC O9ANK5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ID149.
GN ID149.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=110SPC4;
RX MEDLINE=21101824; Pubmed=11157954;
RA Gottfert M., Rotlisberger S., Kundig C., Beck C., Marty R.,
RA Hennecke H.;
RT "Potential symbiosis-specific genes uncovered by sequencing a 410-Kb
RT DNA region of the Bradyrhizobium japonicum chromosome.";
RL J. Bacteriol. 183:1405-1412(2001).
DR EMBL; AF322012; AAG60765.1; -
SQ SEQUENCE 92 AA; 10033 MW; CC40382655E98674 CRC64;

Query Match          100.0%; Score 21; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
   |||
Db 66 AAPF 69

RESULT 40
O95470 PRELIMINARY; PRT; 92 AA.
ID O95470
AC O95470;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MHC class II DQ-beta chain (Fragment).
GN DOB.
OS Bos indicus (Zebu).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KENYA BORAN; TISSUE=BLOOD;
```

RX MEDLINE=96043218; PubMed=7486255;
RA Mareello K.L., Gallagher A., McKeever D.J., Spooner R.L., Russell G.C.;
RT "Expression of multiple DOB genes in Bos indicus cattle.";
RL Anim. Genet. 26:345-349(1995).
DR EMBL: X79348; CAA55903.1; -.
DR InterPro: IPR000353; MHC_II_beta.
DR Pfam: PF00969; MHC_II_beta; 1.
DR ProDom: PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1
FT NON_TER 92
SQ SEQUENCE 92 AA; 11175 MW; D82F524E976CF19F CRC64;

Query Match 100.0%; Score 21; DB 7; Length 92;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
1111
Db 79 AAPF 82

Search completed: December 6, 2002, 13:30:08
Job time : 19.5556 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:19:54 ; Search time 26 Seconds
(without alignments)
20.500 Million cell updates/sec

Title: US-10-033-526-3
Perfect score: 19
Sequence: 1 AAPL 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 19 | 100.0 | 4 | 15 | AAR46220 | Serine protease in |
| 2 | 19 | 100.0 | 4 | 18 | AAW52605 | Serine protease-in |
| 3 | 19 | 100.0 | 4 | 18 | AAW29410 | Elastolytic protei |
| 4 | 19 | 100.0 | 4 | 18 | AAW26080 | Substrate #3 for s |
| 5 | 19 | 100.0 | 4 | 18 | AAW12811 | Synthetic substrat |
| 6 | 19 | 100.0 | 4 | 19 | AAW51756 | Substrate used to |
| 7 | 19 | 100.0 | 4 | 19 | AAW37769 | Amino acid sequenc |
| 8 | 19 | 100.0 | 4 | 19 | AAW39105 | Calf thymus immuno |
| 9 | 19 | 100.0 | 4 | 20 | AAV31241 | Alpha-aminoalkyl p |
| 10 | 19 | 100.0 | 4 | 21 | AAB03091 | Substrate peptide |

| | | | | | | |
|----|----|-------|----|----|----------|--------------------|
| 11 | 19 | 100.0 | 4 | 22 | AAB97637 | Alpha-ketoamide el |
| 12 | 19 | 100.0 | 4 | 22 | AAB35964 | Elastase cleavable |
| 13 | 19 | 100.0 | 4 | 23 | AAO15231 | Porphyromonas ging |
| 14 | 19 | 100.0 | 4 | 23 | AAO18045 | C-terminal truncat |
| 15 | 19 | 100.0 | 4 | 23 | ABB83183 | Synthetic peptide |
| 16 | 19 | 100.0 | 8 | 16 | AAR69616 | MHC class I-deri |
| 17 | 19 | 100.0 | 8 | 16 | AAR69617 | MHC class I-deri |
| 18 | 19 | 100.0 | 9 | 17 | AAW07118 | Synthetic peptide |
| 19 | 19 | 100.0 | 9 | 20 | AAV46545 | Immunogenic peptid |
| 20 | 19 | 100.0 | 9 | 20 | AAV46769 | Immunogenic peptid |
| 21 | 19 | 100.0 | 9 | 20 | AAV46771 | Immunogenic peptid |
| 22 | 19 | 100.0 | 9 | 20 | AAV47521 | Immunogenic peptid |
| 23 | 19 | 100.0 | 9 | 20 | AAV47770 | Immunogenic peptid |
| 24 | 19 | 100.0 | 9 | 21 | AAB48605 | C-Abl Tyr kinase a |
| 25 | 19 | 100.0 | 9 | 22 | AAU06954 | Human MHC molecule |
| 26 | 19 | 100.0 | 9 | 22 | AAU02273 | HLA binding TADG-1 |
| 27 | 19 | 100.0 | 10 | 17 | AAR96502 | Hepatitis C virus |
| 28 | 19 | 100.0 | 10 | 18 | AAW16482 | Acceptor peptide f |
| 29 | 19 | 100.0 | 10 | 20 | AAV46558 | Immunogenic peptid |
| 30 | 19 | 100.0 | 10 | 22 | AAU06981 | Human MHC molecule |
| 31 | 19 | 100.0 | 10 | 22 | AAU06981 | Human MHC molecule |
| 32 | 19 | 100.0 | 12 | 14 | AAR42595 | Peptide Jb1 analog |
| 33 | 19 | 100.0 | 12 | 19 | AAW51870 | Peptide sequence w |
| 34 | 19 | 100.0 | 12 | 22 | AAB92048 | Growth factor pept |
| 35 | 19 | 100.0 | 13 | 21 | AAB09773 | Insulin like growt |
| 36 | 19 | 100.0 | 14 | 14 | AAR38723 | N-terminal fragmen |
| 37 | 19 | 100.0 | 14 | 15 | AAR53987 | Meg-Pot N-terminal |
| 38 | 19 | 100.0 | 14 | 22 | ABB56859 | Human SNP related |
| 39 | 19 | 100.0 | 14 | 22 | AAB67038 | Human insulin-like |
| 40 | 19 | 100.0 | 14 | 23 | ABP46386 | Human Blys binding |
| 41 | 19 | 100.0 | 14 | 23 | ABP46390 | Human Blys binding |
| 42 | 19 | 100.0 | 14 | 23 | ABP46392 | Human Blys binding |
| 43 | 19 | 100.0 | 14 | 23 | ABP46394 | Human Blys binding |
| 44 | 19 | 100.0 | 14 | 23 | ABP46396 | Human Blys binding |
| 45 | 19 | 100.0 | 14 | 23 | ABB57702 | IGFBP-1 binding pe |

ALIGNMENTS

RESULT 1
AAR46220
ID AAR46220 standard; peptide; 4 AA.

AC AAR46220;
DT 04-AUG-1994 (first entry)
DE Serine protease inhibitor tetrapeptide.
XX
XX
KW Prevention; schistosomiasis; parasite; infection; prevention;
KW parasitic penetration; skin; cercariae; anti-penetrant.
XX
XX
OS Synthetic.

XX
XX
FH Key
FT Modified-site 1 Location/Qualifiers
FT Modified-site 4 /note= "Bg(peptide blocking gp.) attached"
FT FT /note= "Pl(protease inhibitor), other than
FT chloromethyl ketone, attached"

PN US5284829-A.
XX
XX 08-FEB-1994.
PD
XX 26-NOV-1991; 91US-0798565.
XX
XX 26-NOV-1991; 91US-0798565.
PR
XX
XX (REGC) UNIV CALIFORNIA.
PA
XX
XX Cohen FE, McKerrow JH;
PI

XX
DR WPI; 1994-056364/07.
XX
PT Synthetic tetra:peptide(s) with an N-terminal blocking gp. and
PT C-terminal enzyme inhibitor - can be formulated into soaps and
PT sprays and used to prevent schistosomal skin penetration
XX
PS Disclosure; Page 7; 35pp; English.
XX
CC The sequence is that of a synthetic tetrapeptide serine protease
CC inhibitor which can be used to prevent schistosome parasite
CC infection. It may be used in a formulation as a soap, lotion,
CC cream, spray, etc. to stop parasitic penetration of the skin.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 15; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 1 AAPL 4

RESULT 2
ID AAW52605 standard; peptide; 4 AA.
XX
AC AAW52605;
XX
DT 22-JUN-1998 (first entry)
XX
DE Serine protease-inhibiting peptide with C-terminal phosphonate residue.
XX
KW Selective; serine protease inhibitor; trypsin; elastase; chymotrypsin;
KW antiinflammatory; anticoagulant; antitumour; Schistosoma mansoni.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site 4 /note= "MeO-Suc-Ala"
FT Modified-site 4 /note= "Leu(P)-(Oph)2; where Leu(P) represents a
FT leucine analogue in which the carboxyl group
FT -COOH has been replaced by a phosphonate group
FT -P(=O)(OH)(OH); and (Oph)2 indicates that the
FT phosphonate has been diphenyl esterified"
XX
PN USS686419-A.
XX
PD 11-NOV-1997.
XX
PF 21-JAN-1994; 94US-0184286.
XX
PR 21-JAN-1994; 94US-0184286.
XX
PA (GEOR-) GEORGIA TECH RES CORP.
PI Boduszek B, Oleksyszyn J, Powers JC;
XX WPI; 1997-558177/51.
DR
XX
PT New peptidyl derivatives of alpha-aminoalkyl phosphonic acid
PT di:ester - are serine protease inhibitors, useful e.g. for reducing
PT blood coagulation, controlling tumour invasion or treating
PT inflammation
XX
PS Example 13; Column 22; 16pp; English.
XX
CC The patent discloses new peptidyl derivatives of diesters of alpha-
CC aminoalkylphosphonic acids having basic substituents, of formula

CC X-AA4-AA3-AA2-NH-CHR-P(=O)(OZ)(OZ1), in which: R = B-substituted
CC phenyl, B-substituted benzyl or B-substituted 1-6C alkyl; B = amidino,
CC guanidino, isothioureido or amino; Z and Z1 = 1-6C perfluoroalkyl or
CC optionally substituted phenyl; AA2 = an L- or D-amino acid residue;
CC AA3 = a single bond or an L- or D-amino acid residue; AA4 = a single
CC bond or an L- or D-amino acid residue; X = H, NH2CO, NH2CS, NH2SO2,
CC YNHCO, YNHCS, YNHSO2, YCS, YSO2, YOCO, YOCS or YCO; Y = optionally
CC naphthyl, or 1-6C alkyl with 1-2, optionally substituted, attached
CC phenyl groups. The new compounds are inhibitors of trypsin, elastase,
CC chymotrypsin and other serine proteases; and their inhibitory activity
CC is selective depending on the identity of the alpha-aminoalkylphosphonic
CC acid ester residue. They can be used as antiinflammatory agents,
CC anticoagulants and antitumour agents. The present sequence is
CC an example of a compound which can specifically inhibit the activity
CC of elastolytic proteinase from Schistosoma mansoni.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 18; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 1 AAPL 4

RESULT 3
ID AAW29410 standard; peptide; 4 AA.
XX
AC AAW29410;
XX
DT 23-FEB-1998 (first entry)
XX
DE Elastolytic proteinase inhibitor peptide ketoamide derivative.
XX
KW Peptide ketoamide derivative; elastolytic proteinase inhibitor;
KW Alzheimer's disease; coagulation disorder; serine protease;
KW cysteine protease; calpain; cathepsin G; neurodegenerative disease;
KW ischaemia; stroke; trypsin inhibitor; chymase inhibitor; anticoagulant;
KW tissue damage; thrombosis; blistering; Schistosoma mansoni.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site 4 /note= "Succinyl-ala"
FT Modified-site 4 /note= "Leu-NR3R4, where R3 and R4 are selected
FT independently from the group consisting of H,
FT 1-20C alkyl, 3-20C cyclised alkyl, 1-20C
FT alkyl with a phenyl group (optionally mono-,
FT di- or tri-substituted with K) attached to the
FT 1-20C alkyl, 3-20C cyclised alkyl with an
FT attached phenyl group substituted with K, 1-20
FT alkyl with a morpholine, a pyrrolidine or a
FT piperidine ring attached through nitrogen to the
FT alkyl, 1-20C alkyl with an OH group attached to
FT the alkyl, 1-10C with an attached pyridyl
FT group or cyclohexyl group, -NCH2CH2-
FT (4-hydroxyphenyl)- and -NHCH2CH2-(3-indolyl)"
XX
PN USS610297-A.
XX
PD 11-MAR-1997.
XX
PF 06-OCT-1995; 95US-0815073.
XX
PR 27-DEC-1991; 91US-0815073.
PR 09-SEP-1993; 93US-0118997.
PR 20-MAY-1994; 94US-0247081.

XX
PA (GEOR-) GEORGIA TECH RES CORP.
XX
PI Powers JC;
XX
DR WPI; 1997-178454/16.
XX
PT New peptide keto:amide derivs - are protease inhibitors used to
PT treat e.g. Alzheimer's disease, coagulation disorders and other
PT neurodegenerative disorders.
XX
PS Disclosure; Columns 17-18; 17pp; English.
XX
CC The present sequence represents a peptide ketoamide derivative which is
CC useful for selectively inhibiting elastolytic proteinase from
CC Schistosoma mansoni. It is a specific example of a new class of peptides
CC which selectively inhibit serine proteases or cysteine proteases,
CC including calpains and cathepsin B. The calpain inhibitors are useful for
CC treatment of various neurodegenerative diseases and conditions including
CC ischaemia, stroke and Alzheimer's disease. The protease inhibitors,
CC especially the elastase, trypsin and chymase inhibitors, are used to
CC control tissue damage and various inflammatory conditions mediated by
CC proteases, such as blistering. They are also useful as anticoagulants and
CC can be used to treat thrombosis. The peptides may also be used to
CC identify new proteolytic enzymes encountered in research. Further, they
CC may also be useful in research and industrially to prevent undesired
CC proteolysis that occurs during the production, isolation, purification,
CC transport and storage of valuable peptides and proteins; e.g. they may be
CC added to antibodies, enzymes, plasma proteins, tissue extracts or other
CC proteins and peptides which are widely sold for use in clinical analyses,
CC biomedical research and for many other reasons.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 18; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
| | | |
Db 1 AAPL 4

RESULT 4
AAW26080
ID AAW26080 standard; peptide; 4 AA.
XX
AC AAW26080;
XX
DT 05-NOV-1997 (first entry)
XX
DE Substrate #3 for serine protease.
XX
KW Serine protease; substrate; Streptomyces griseus; guanidine; pre-soak;
KW cleaning composition; laundry detergent; additive composition; enzyme;
KW dishwasher detergent; drain opener; urea; contact lens cleanser;
KW proteinaceous stain.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site /note= "succinylated"
FT Modified-site 4
FT /note= "para-nitroanilidated"
XX
PN US5646028-A.
XX
PD 08-JUL-1997.
XX
PF 18-JUN-1991; 91US-0718303.
XX
PR 18-JUN-1991; 91US-0718303.

PR 06-NOV-1992; 92US-0973343.
PR 18-AUG-1994; 94US-0292924.
PR 17-OCT-1995; 95US-0544143.
XX
PA (CLR) CLOROX CO.
XX
PI Leigh SD;
XX
DR WPI; 1997-362936/33.
XX
PT Serine protease from Streptomyces griseus ATCC 55178 - with good
PT stability in presence of urea or guanidine, useful in cleaning
PT compositions, including laundry and dishwashing detergents
XX
PS Example 1; Column 12; 16pp; English.
XX
CC AAW26078-W26096 represent substrates for the serine protease of the
CC invention. The serine protease recognises these sequences, but is
CC specific for the sequence shown in AAW24567. The protease has the
CC N-terminal and C-terminal sequences represented by AAW24565 and AAW24566
CC respectively. The serine protease was isolated from Streptomyces griseus
CC variety alkaphilus No. 33 (ATCC 55178). The protease has an apparent
CC molecular weight of 19 kD (by reducing sodium dodecylsulphate
CC polyacrylamide gel electrophoresis), and improved stability against urea
CC and guanidine. The protease is inhibited by phenylmethylsulphonyl
CC fluoride. The serine protease is useful in liquid or granular cleaning
CC compositions, specifically laundry detergents or additive compositions.
CC It is also useful in automatic dishwasher detergents, pre-soaks, drain
CC openers, contact lens cleansers etc. The protease has better activity
CC against proteinaceous stains than known enzymes and unusually high
CC stability in the presence of chaotropic agents.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 18; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
| | | |
Db 1 AAPL 4

RESULT 5
AAW12811
ID AAW12811 standard; peptide; 4 AA.
XX
AC AAW12811;
XX
DT 21-APR-1997 (first entry)
XX
DE Synthetic substrate #2 for cold-adapted alkali protease.
XX
KW Cold-adapted alkali protease; alteromonas; urea-denatured haemoglobin;
KW yolk; casein; p-nitroaniline; detergent; meat softener.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site /label= Succinylated
FT Modified-site 4
FT /label= p-nitroanilnylated
XX
PN JP08322562-A.
XX
PD 10-DEC-1996.
XX
PF 01-JUN-1995; 95JP-0135056.
XX
PR 01-JUN-1995; 95JP-0135056.
XX
PA (KAOS) KAO CORP.

XX
DR WPI; 1997-081078/08.
XX
PT Cold-adapted alkali protease - produced by Alteromonas, used in
PT detergents
XX
PS Claim 1; Page 2; 10pp; Japanese.
XX
CC AAW12810-W12812 represent synthetic substrates for the cold-adapted
CC alkali protease of the invention. The alkali protease of the invention
CC is produced by Alteromonas, specifically by Alteromonas species KSM-SP
CC 111 which is isolated from a shellfish in the Antarctic ocean. The
CC enzyme of the invention has an active temperature of 0-70 degrees C, with
CC an optimum of 40 degrees, and with 20% of activity retained at 10
CC degrees, and 10% maintained at 0 degrees. The pH range of the enzyme is
CC 4-14, with an optimum of 11, but 70% (or higher) of the activity
CC maintained at pH12. The protease has a molecular weight of 54000, as
CC determined by SDS-polyacrylamide gel electrophoresis. The enzyme is
CC capable of acting on casein, urea-denatured haemoglobin, yolk, and the
CC synthetic substrates represented by these sequences, to yield
CC p-nitroaniline. The enzyme is inhibited by the Hg and Zn metal ions,
CC EDTA, phenylmethanesulphonyl fluoride, chymostatin or
CC p-chloromercuribenzoic acid. Activity of the enzyme is increased 2-7
CC times by the presence of sodium alkanesulphate, sodium
CC alpha-olefinesulphate, sodium polyoxyethylenealkylsulphate, softanol 70H
CC or alpha-sulphofatty acid ester. The protease can be used as a component
CC of detergents, or as a softener for meat when used at a lower
CC temperature.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 18; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 6
AAW51756
ID AAW51756 standard; peptide; 4 AA.
XX
AC AAW51756;
XX
DT 10-SEP-1998 (first entry)
XX
DE Substrate used to assay rotamase activity of 50-54 kDa immunophilin.
XX
KW immunophilin; tryptic peptide; C-terminus; FK-506; cyclosporine A;
KW rapamycin; immunosuppressant drug; binding assay.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "Succinyl-Ala"
FT Modified-site 4 /note= "Leu-para-nitroanilide"
FT
FT
XX
PN US5780307-A.
XX
PD 14-JUL-1998.
XX
PF 26-JUL-1996; 96US-0686759.
XX
XX 23-FEB-1994; 94US-0200404.
PR 02-DEC-1988; 88US-0279176.
PR 02-MAR-1990; 90US-0487115.
PR 09-MAY-1990; 90US-0521074.
PR 22-OCT-1991; 91US-0782761.
PR 26-FEB-1992; 92US-0841792.

PR 26-JUL-1996; 96US-0686759.
XX
PA (SOLD/) SOLDIN S J.
XX
PI Soldin SJ;
XX
DR WPI; 1998-413070/35.
XX
PT New immunophilin from mammalian lymphoid tissue - used in binding
PT assays for cyclosporine, FK-506 and rapamycin, and for recovering new
PT immunosuppressants from extracts and cell cultures
XX
PS Example 18; Columns 31-32; 56pp; English.
XX
CC The invention relates to an immunophilin (i.e. a cytosolic immuno-
CC suppressant drug binding protein) that binds specifically to FK-506,
CC cyclosporine A and rapamycin; has molecular weight 34-47 kD (by
CC chromatography, sodium dodecylsulphate-polyacrylamide gel
CC electrophoresis or amino acid analysis) and has isoelectric point (pI)
CC 6.5-7. It is optionally immobilised on a solid support, is used in
CC binding assays for the specified immunosuppressants, and for capturing
CC potential immunosuppressants from microbial extracts or cell cultures,
CC e.g. active metabolites of the specified immunosuppressants or their
CC mammalian homologues. The present sequence represents a peptide
CC substrate which was used to assay the rotamase activity of a 50-54 kDa
CC immunophilin.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 7
AAW37769
ID AAW37769 standard; Protein; 4 AA.
XX
AC AAW37769;
XX
DT 20-JUL-1998 (first entry)
XX
DE Amino acid sequence of synthetic chromogenic substrate.
XX
KW Subtilisin-type serine protease inhibitor; inhibition; ss;
KW proteolytic activity; extracellular elastolytic serine protease;
KW Aspergillus hyphae; lung; germination; aspergillosis; nasal delivery;
KW chromogenic substrate.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Active-site 1
FT Modified-site 1 /note= "Suc-Ala"
FT Modified-site 4 /note= "Leu-pNa"
FT
FT
XX
PN US5739283-A.
XX
PD 14-APR-1998.
XX
PF 07-JUN-1995; 95US-0486895.
XX
XX 07-JUN-1995; 95US-0486895.
PR
XX (OHIO-) OHIO STATE RES FOUND.
PA
XX
PI Ceselski SK, Copelan EA, Kolattukudy PE, Markaryan AN;

```
XX
DR WPI, 1998-250492/22.
XX
PT Treatment of invasive pulmonary aspergillosis - by nasal
PT administration of composition comprising subtilisin type serine
PT protease inhibitor e.g. streptomycetes subtilisin inhibitor, and
XX carrier e.g. phosphate buffered saline
XX
PS Claim 6; Columns 17-18; 12pp; English.
XX
CC The amino acid sequence is a chromogenic substrate used in the
CC method of invention to develop subtilisin-type serine protease
CC inhibitor (SSI). The SSI inhibits the proteolytic activity of the
CC extracellular elastolytic serine protease produced by aspergillus and
CC reduces the invasion of the lung and tissues surrounding it by the
CC germinating hyphae of aspergillus. It can be used in a drug
CC composition to treat aspergillosis via nasal delivery.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
   ||||
Db 1 AAPL 4

RESULT 8
AAW39105
ID AAW39105 standard; peptide; 4 AA.
XX
AC AAW39105;
XX
DT 08-APR-1998 (first entry)
XX
DE Calf thymus immunophilin rotamase substrate 6.
XX
KW Immunophilin; calf; immunosuppressant drug; FK506; rapamycin;
KW cyclosporin A; rotamase; therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site /note= "N-terminal Suc-Ala"
FT Modified-site 4 /note= "C-terminal Leu-PNA"
XX
PN US5698448-A.
XX
PD 16-DEC-1997.
XX
PE 08-APR-1994; 94US-0224868.
XX
PR 23-FEB-1994; 94US-0200404.
PR 02-DEC-1988; 88US-0279176.
PR 02-MAR-1990; 90US-0487115.
PR 03-MAY-1990; 90US-0521074.
PR 22-OCT-1991; 91US-0782761.
PR 26-FEB-1992; 92US-0841792.
PR 08-APR-1994; 94US-0224868.
XX
PA (SOLD/) SOLDIN S J.
XX
PI Soldin SJ;
XX
DR WPI, 1998-051496/05.
XX
PT Immunophilin protein that binds immunosuppressant drugs - useful in
PT assays for cyclosporin A, FK506 and rapamycin
XX
```

```
PS Example 18; Column 31; 54pp; English.
XX
CC Peptides AAW39100-W39110 represent substrates used in an assay to
CC investigate the rotamase activity of a 10-12 kDa and a 50-52 kDa
CC immunophilin isolated from calf thymus. Immunophilins are proteins
CC capable of binding to immunosuppressant drugs, in this example FK506,
CC rapamycin or cyclosporin. This substrate has a rotamase activity of
CC 0.45 R-obs (min-1) with the 50-52 kDa immunophilin and 0.93 K-obs
CC (min-1) with the 10-12 kDa immunophilin. Immunophilins will have
CC applications for monitoring immunosuppressive therapy and for the capture
CC of potential immunosuppressive drugs from microbial extract, culture
CC media or from mammalian body fluids and tissues.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
   ||||
Db 1 AAPL 4

RESULT 9
AAV31241
ID AAV31241 standard; peptide; 4 AA.
XX
AC AAV31241;
XX
DT 09-NOV-1999 (first entry)
XX
DE Alpha-aminoalkyl phosphonate derivative peptide 3.
XX
KW Alpha-aminoalkyl phosphonate; serine protease inhibitor; anticoagulant;
KW cytosstatic; anti-inflammatory; gastrointestinal; endocrine; respiratory;
KW dermatological; chymotrypsin inhibitor; elastase inhibitor; inflammation;
KW trypsinase inhibitor; trypsin inhibitor; blood coagulation; pancreatitis;
KW tumour invasion control; emphysema; respiratory distress syndrome;
KW skin blistering.
XX
OS Synthetic.
XX
PN US5952307-A.
XX
PD 14-SEP-1999.
XX
PE 14-AUG-1997; 97US-0907840.
XX
PR 14-AUG-1997; 97US-0907840.
PR 21-JAN-1994; 94US-0184286.
XX
PA (GEOR-) GEORGIA TECH RES CORP.
XX
PI Jackson DS, Ni L, Powers JC;
XX
DR WPI, 1999-539191/45.
XX
PT New aminoalkyl phosphonate derivatives are serine protease
PT inhibitors useful for decreasing blood coagulation and for treating
PT tumour invasion, inflammation, pancreatitis, emphysema, respiratory
PT distress syndrome and skin blistering
XX
PS Disclosure; Column 31-32; 18pp; English.
XX
CC This invention describes the production of novel alpha-aminoalkyl
CC phosphonate derivatives (I) which have anticoagulant, cytosstatic,
CC anti-inflammatory, gastrointestinal, endocrine, respiratory and
CC dermatological activity. The products of the invention are also found
CC to act as serine protease inhibitors, chymotrypsin inhibitors, elastase
CC inhibitors, trypsinase inhibitors and trypsin inhibitors. (I) are useful
CC for decreasing blood coagulation, and for controlling tumour invasion,
CC inflammations, pancreatitis, emphysema, respiratory distress syndrome
```

```
CC and skin blistering. AAY31239-V31241 are peptides used to describe the
CC method of the invention.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 1 AAPL 4

RESULT 10
AAB03091
ID AAB03091 standard; peptide; 4 AA.
XX
AC AAB03091;
XX
DT 10-OCT-2000 (first entry)
XX
DE Substrate peptide #2.
XX
KW Serine protease; trypsin activity; Trichoderma; bacterial; coagulant;
KW hypertensive; antiinflammatory; leather preparation; silk treatment.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "Conjugated to succinyl moiety"
FT Modified-site 4 /note= "Conjugated to p-nitroaniline (pNA)"
XX
PN Jp2000116377-A.
XX
PD 25-APR-2000.
XX
PF 08-OCT-1998; 98JP-0303263.
XX
PR 08-OCT-1998; 98JP-0303263.
XX
PA (AMANO ) AMANO PHARM KK.
XX
DR WPI; 2000-369402/32.
XX
PT A new serine protease and its preparation, used clinically in blood
PT coagulation, hypotension and anti-inflammation -
XX
PS Example 3; Page 5; 9pp; Japanese.
XX
CC The invention relates to a novel serine protease from Trichoderma sp.
CC No.9064. The N-terminus of this protease is given in AAB03085. The novel
CC protease has trypsin-like activity, specifically cleaving peptide chains
CC on the carboxyl side of a basic amino acid (e.g., arginine or lysine).
CC It has a pH optimum of 7-8, and is stable between pH 5 and pH10. It has
CC a temperature optimum of approximately 40 degrees Celsius, and is stable
CC between 40 and 50 degrees Celsius. The invention also relates to a
CC method for the preparation of the novel serine protease, and the use of
CC the protease in protein degradation. The enzyme has coagulant,
CC hypertensive and anti-inflammatory effects. It may also be used in the
CC preparation of leather, for raw silk treatment and for the preparation of
CC protein hydrolysate. Sequences AAB03090-B03094 represent tetrapeptides
CC used in the determination of the activity of the novel protease.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 21; Length 4;
Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
```

```
Db 1 AAPL 4

RESULT 11
AAB97637
ID AAB97637 standard; peptide; 4 AA.
XX
AC AAB97637;
XX
DT 21-SEP-2001 (first entry)
XX
DE Alpha-ketoamide elastolytic proteinase inhibitor peptide.
XX
KW Peptide ketoamide; serine protease inhibitor;
KW cysteine protease inhibitor; transition state analogue;
KW neurodegenerative disease; ischaemia; stroke; Alzheimer's disease;
KW anticoagulant; thrombosis; adult respiratory distress syndrome;
KW emphysema; rheumatoid arthritis; pancreatitis; viral infection;
KW muscular dystrophy; myocardial tissue damage; tumour metastasis;
KW bone resorption; Schistosoma mansoni; elastolytic protease inhibitor.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "Suc-Ala"
FT Modified-site 4 /note= "Leu-CO-NR3R4"
XX
PN US6235929-B1.
XX
PD 22-MAY-2001.
XX
PF 27-DEC-1996; 96US-0777354.
XX
PR 27-DEC-1991; 91US-0815073.
PR 09-SEP-1993; 93US-0118997.
PR 20-MAY-1994; 94US-0246511.
PR 06-OCT-1995; 95US-0539944.
XX
PA (GEOR-) GEORGIA TECH RES CORP.
XX
PI Powers JC;
XX
DR WPI; 2001-440210/47.
XX
PT New tripeptide ketoamide derivatives are serine and cysteine protease
PT inhibitors, useful as anticoagulants and for treating neurodegenerative
PT diseases, thrombosis, emphysema and rheumatoid arthritis -
XX
PS Disclosure; Column 18; 24pp; English.
XX
CC The invention relates to peptide ketoamide derivatives useful for
CC selectively inhibiting serine proteases and cysteine proteases, and for
CC generally inhibiting all members of these classes of enzymes. The peptide
CC ketoamides are derivatives of substrate peptides of serine and cysteine
CC proteases and act as transition state analogues. The peptides of the
CC invention have the formula M1-AA1-AA2-AA3-CO-NR3R4 where:
CC M1 is NH2CO-, NH2CS-, NH2SO2-, XNHCO-, X2NCO-, XNHCS-, X2NCS-,
CC XNHSO2-, X2NSO2-, XCO-, XCS-, XSO2-, XOCO- or XOCS;
CC X is 1-10C (fluoro)alkyl (optionally substituted by J), 1-adamantyl,
CC 9-fluorenyl, phenyl or naphthyl (optionally substituted by up to 3 of
CC K), or 1-10C alkylphenyl, 1-10C alkylidiphenyl or 1-10C alkylphenoxy
CC (all optionally substituted by K);
CC J is halo, COOH, OH, CN, NO2, NH2, 1-10C alkoxy, 1-10C alkylamine,
CC 2-12C dialkylamine, 1-10C alkyl-OCO-, 1-10C alkyl-OC(=NH)- or 1-10C
CC alkylthio;
CC K is halo, 1-10C (perfluoro)alkyl, 1-10C alkoxy, NO2, CN, OH, COOH,
CC NH2, 1-10C alkylamino, 2-12C dialkylamino, 1-10C acyl, 1-10C alkoxy-CO
CC or 1-10C alkylthio;
CC AA1, AA2 are independently Ala, Val, Leu, Ile, Gly, Ser, Asp or Glu;
CC AA3 is Asp or Glu, in either the L or D form;
```

CC R3 is 2-3C alkylphenyl, 3-20C cycloalkylphenyl, 1-20C alkylphenyl
CC (substituted by up to 3 of K), 3-20C cycloalkylphenyl (substituted by
CC K), NH-CH2CH2-(4-hydroxyphenyl) or NH-CH2CH2-(3-indolyl); and
CC R4 = H, 3-20C alkyl, cycloalkyl, 1-20C alkylphenyl (optionally
CC substituted by up to 3 of K), 3-20C cycloalkylphenyl (optionally
CC substituted by K), NHCH2CH2-(4-hydroxyphenyl) or
CC NHCH2CH2-2-(3-indolyl).
CC The peptide ketoamides are useful for treating neurodegenerative diseases
CC (including ischaemia, stroke and Alzheimer's disease), as anticoagulants,
CC and for treating thrombosis. They are also useful for treating
CC emphysema, adult respiratory distress syndrome, rheumatoid arthritis,
CC pancreatitis, viral infections, muscular dystrophy, myocardial tissue
CC damage, tumour metastasis and bone resorption. The present sequence
CC represents a Schistosoma mansoni elastolytic protease peptide
CC inhibitor.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 22; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 1 AAPL 4

RESULT 12
AAB35964
ID AAB35964 standard; peptide: 4 AA.
XX
AC AAB35964;
XX
DT 01-MAR-2001 (first entry)
XX
DE Elastase cleavable peptide #2.
XX
KW Polymeric drug conjugate; enzymatically cleavable linker; cardiant;
KW antiinflammatory; cytostatic; hepatotropic; neuroprotective; cancer;
KW antibacterial; nephrotropic; immunomodulatory; neoplastic disease;
KW chronic inflammatory disease; acute inflammatory disease;
KW cardiac disease; renal disease; liver disease; lung disease;
KW neurological disease; musculoskeletal disease; immunological disorder;
KW microbial infection.
XX
OS Synthetic.
XX
PN WO200064486-A2.
XX
PD 02-NOV-2000.
XX
PF 28-APR-2000; 2000WO-US11670.
XX
XX 28-APR-1999; 99US-0131404.
PR 02-NOV-1999; 99US-0163090.
XX
XX
PA (VERI-) VERITAS MEDICAL TECHNOLOGIES INC.
XX
PI Pachence JM, Belinka BA, Ramani T;
XX
DR WPI; 2001-031659/04.
XX
XX Polymeric drug conjugate, for treating diseases associated with organs
PT e.g. liver or heart, has biologically active agents linked to regular
PT repeating linear or branched co-polymers by enzymatically cleavable
PT marker -
XX
XX Claim 30; Page 93; 100pp; English.
XX
XX This invention relates to a polymeric drug conjugate. The drug conjugate
CC comprises biologically active agents conjugated via an enzymatically
CC cleavable linker to a regular repeating linear unit comprising a water
CC soluble polymer segment and a multifunctional chemical moiety, or to a

CC branched polymer comprising two or more water soluble polymer segments
CC each bound to a common multifunctional chemical moiety. The polymeric
CC drug conjugate has antiinflammatory; cytostatic; cardiant; hepatotropic;
CC neuroprotective; antibacterial; nephrotropic; and immunomodulatory
CC activity. The drug conjugate is useful for alleviating a pathological
CC condition such as neoplastic diseases, chronic inflammatory diseases
CC acute inflammatory diseases, cardiac diseases, renal diseases, liver
CC diseases, lung diseases, neurological diseases, musculoskeletal diseases
CC and immunological disorders and various microbial infections by
CC modulating immunological or hormonal function. The present sequence
CC represents a peptide which can be used as the enzymatically cleavable
CC linker in the drug conjugate of the invention.
XX
SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 22; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 1 AAPL 4

RESULT 13
AAO15231
ID AAO15231 standard; Peptide: 4 AA.
XX
AC AAO15231;
XX
DT 05-SEP-2002 (first entry)
XX
DE Porphyromonas gingivalis dipeptidylpeptidase-7 substrate peptide 10.
XX
XX Dipeptidylpeptidase-7; DPP-7; amidolytic cleavage; substrate peptide;
KW DPP-7 inhibitor identification; periodontal disease; gingivitis;
KW periodontitis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "The residue is modified with N-succinyl"
FT Modified-site 4 /note= "The residue is modified with PNA"
FT
XX
PN WO200238742-A2.
XX
PD 16-MAY-2002.
XX
PF 08-NOV-2001; 2001WO-US46782.
XX
XX 08-NOV-2000; 2000US-246827P.
PR
XX (UYGE-) UNIV GEORGIA RES FOUND INC.
XX
XX Travis J, Potempa JS, Banbula A, Bugno M;
XX
DR WPI; 2002-490075/52.
XX
XX Novel isolated dipeptidylpeptidase useful for identifying inhibitor of
PT the dipeptidylpeptidase for protecting an animal from periodontal
PT disease caused by Porphyromonas gingivalis -
XX
XX Example 5; Page 29; 65pp; English.
XX
XX The invention comprises the amino acid and coding sequence of the
CC Porphyromonas gingivalis dipeptidylpeptidase-7 (DPP-7) enzyme. The DPP-7
CC enzyme has amidolytic activity for cleavage of a peptide bond between the
CC second and third amino acids from the N-terminal end of a target peptide.
CC The DPP-7 target peptide has an aliphatic or aromatic residue as a
CC substituent on the alpha-carbon atom of the second amino acid from the
CC N-terminal end. The DPP-1 protein and DNA sequences of the invention are

CC useful for identifying a DPP-7 inhibitor. DPP-7 inhibitors are useful for
CC reducing the growth of a bacterium and protecting an animal from a
CC periodontal disease caused by Porphyromonas gingivalis (e.g. gingivitis
CC or periodontitis). The present amino acid sequence represents a substrate
CC peptide that was not hydrolysed by the Porphyromonas gingivalis
CC dipeptidylpeptidase-7 (DPP-7) enzyme of the invention.

XX SQ. Sequence 4 AA;

Query Match 100.0%; Score 19; DB 23; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 14

AAO18045 ID AAO18045 standard; peptide; 4 AA.

XX AC AAO18045;

XX DT 30-AUG-2002 (first entry)

XX DE C-terminal truncated apoE formation inhibitor peptide #3.

KW C-terminal truncated apoE; apolipoprotein E; apoE; Alzheimer's disease;
KW neurofibrillary tangle; apoE4 allele; nootropic; neuroprotective;
KW cardiant; vulnery; cerebroprotective; coronary artery disease;
KW head trauma; stroke.

XX OS Synthetic.

XX PN WO200238108-A2.

XX PD 16-MAY-2002.

XX PF 02-NOV-2001; 2001WO-US51172.

XX PR 03-NOV-2000; 2000US-245737P.

XX PA (GLAD-) GLADSTONE INST J DAVID.

XX PI Huang Y, Mahley RW;

XX DR WPI; 2002-490051/52.

XX PT Inhibiting neurofibrillary tangles formation, useful for treating e.g.
PT Alzheimer's, coronary artery disease or stroke, by reducing the
PT formation of carboxyl-terminal truncated form of apolipoprotein E in a
PT neuron of the individual -

XX PS Claim 24; Page 62; 75pp; English.

XX CC The present invention relates to a method of inhibiting the formation of
CC neurofibrillary tangles in an individual, which involves reducing the
CC formation of a carboxyl-terminal truncated form of apolipoprotein E
CC (apoE) in a neurone in the individual. The method is useful for
CC inhibiting the formation of neurofibrillary tangles in an individual. The
CC reduction in the formation of carboxyl-terminal truncated apoE treats a
CC disorder related to apoE in an individual, specifically Alzheimer's
CC disease, coronary artery disease, head trauma or stroke. The present
CC sequence is a peptide capable of inhibiting the formation of the
CC C-terminal truncated form of apoE.

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 23; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 15

ABB83183 ID ABB83183 standard; peptide; 4 AA.

XX AC ABB83183;

XX DT 16-AUG-2002 (first entry)

XX DE Synthetic peptide #2 used to assay Pral protease activity.

XX KW Pral; protease; enzyme; fungicide; insecticide; serine-peptidase.

XX OS Synthetic.

XX FH Key location/Qualifiers

FT Modified-site 1 /note= "N-terminal succinyl"

FT Modified-site 4 /note= "Leu-pNA"

XX PN WO200244359-A1.

XX PD 06-JUN-2002.

XX PF 30-NOV-2001; 2001WO-ES00471.

XX PR 01-DEC-2000; 2000ES-0002897.

XX PA (NEWB-) NEWBIOTECHNIC SA.

XX PA (UYSE-) UNIV SEVILLA.

XX PA (UYSA-) UNIV SALAMANCA.

XX PI Suarez Fernandez B, Rey Barrera M, Monte Vazquez E;

XX DR WPI; 2002-471830/50.

XX PT New proteolytic enzyme from Trichoderma harzianum, useful e.g. for
PT protecting plants against fungal attack, also related nucleic acid -

XX PS Example 3; Page 26; 51pp; Spanish.

XX CC The present invention relates to Pral protease from Trichoderma
CC harzianum (ABB83181). Pral, a serine-peptidase, catalyses irreversible
CC inactivation of enzymes and proteins essential for pathogenicity of
CC fungi. Pral can be used to degrade proteins and peptides, especially
CC structural components of the cell walls of fungi, insects and arachnids.
CC Particularly, Pral can be used (optionally in combination with chemical
CC fungicides) to protect plants, animals, harvested crops and foods against
CC fungi; to generate protoplasts and yeast extracts; for recovery of
CC manoproteins; in preparation of wine and (grape) juice; for removal of
CC dental plaque; in tooth- and denture- and contact lens-cleaning solutions;
CC to remove biofilms; to treat or clean textiles; as disinfectant and to
CC prevent contamination of analytical samples. The present peptide was used
CC in an example from the invention to assay Pral activity.

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 23; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 16

AAR69616
ID AAR69616 standard; peptide; 8 AA.
XX
AC AAR69616;
XX
DT 29-AUG-1995 (first entry)
XX
DE MHC class I-derived peptide.
XX
KM MHC class I; major histocompatibility complex; insulin receptor;
KW diabetes.
XX
OS Synthetic.
XX
PN US5385888-A.
XX
PD 31-JAN-1995.
XX
PF 20-MAR-1987; 87US-0028241.
XX
PR 20-MAR-1987; 87US-0028241.
PR 14-MAR-1989; 89US-0323565.
PR 01-FEB-1991; 91US-0649471.
PR 03-MAY-1993; 93US-0057184.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Goodenow RS, Olsson L;
XX
DR WPI; 1995-081582/11.
XX
PT Modulating response of cellular insulin receptor to ligand -
PT using peptide deriv. from MHC class I antigen, partic. to
PT potentiate effect of insulin for treating diabetes
XX
PS Disclosure; Column 23; 15pp; English.
XX
CC Response of an insulin receptor (IR) to a ligand is modulated by
CC contacting mammalian cells having IR on the surface with peptides
CC (given in AAR69608-14) derived from MHC class I antigen. Another
CC peptide provided in the specification is given in AAR69616.
CC Peptides may be linked to e.g. antibodies or incorporated into
CC liposomes for administration.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 19; DB 16; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 5 AAPL 8

RESULT 17
AAR69617
ID AAR69617 standard; peptide; 8 AA.
XX
AC AAR69617;
XX
DT 29-AUG-1995 (first entry)
XX
DE MHC class I-derived peptide.
XX
KM MHC class I; major histocompatibility complex; insulin receptor;
KW diabetes.
XX
OS Synthetic.
XX
PN US5385888-A.
XX
PD 31-JAN-1995.

XX
PF 20-MAR-1987; 87US-0028241.
XX
PR 20-MAR-1987; 87US-0028241.
PR 14-MAR-1989; 89US-0323565.
PR 01-FEB-1991; 91US-0649471.
PR 03-MAY-1993; 93US-0057184.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Goodenow RS, Olsson L;
XX
DR WPI; 1995-081582/11.
XX
PT Modulating response of cellular insulin receptor to ligand -
PT using peptide deriv. from MHC class I antigen, partic. to
PT potentiate effect of insulin for treating diabetes
XX
PS Disclosure; Column 25; 15pp; English.
XX
CC Response of an insulin receptor (IR) to a ligand is modulated by
CC contacting mammalian cells having IR on the surface with peptides
CC (given in AAR69608-14) derived from MHC class I antigen. Another
CC peptide provided in the specification is given in AAR69616.
CC Peptides may be linked to e.g. antibodies or incorporated into
CC liposomes for administration.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 19; DB 16; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 5 AAPL 8

RESULT 18
AAW07118
ID AAW07118 standard; peptide; 9 AA.
XX
AC AAW07118;
XX
DT 23-JAN-1997 (first entry)
XX
DE Synthetic peptide used in GalNac-transferase activity SPA.
XX
KW SPA; scintillation proximity assay; antigen; bead coating; capture;
KW antibody; N-acetyl galactosamine transferase; GalNac transferase;
KW activity; enzyme; O-linked glycosylation.
XX
OS Synthetic.
XX
PN WO9615258-A1.
XX
PD 23-MAY-1996.
XX
PF 08-NOV-1995; 95WO-US13483.
XX
PR 16-NOV-1994; 94US-0340283.
XX
PA (UPJO) UPJOHN CO.
XX
PI Elhammer AP;
XX
DR WPI; 1996-268220/27.
XX
PT Scintillation proximity assay for N-acetyl:galactosaminyl activity
PT - esp. for large scale screening of cpds. for their effect on enzyme
PT activity
XX
PS Claim 14; Page 17; 29pp; English.

XX AAW06985-W07180 are antigenic peptides derived from either the
CC product of the human c-myc oncogene or the FLAG peptide (DYKDDK).
CC The peptides are useful for coating beads used in a scintillation
CC proximity assay for N-acetylgalactosamine (GalNAc)-transferase (GNT)
CC activity. The assay involves fewer steps than known assays and is
CC quicker, producing excellent signal-to-noise ratios. The
CC assay is capable of screening large numbers of cpds. for their
CC ability to affect GNT activity and is thus useful for identifying
CC inhibitors and promoters of glycosylation (in partic. O-linked
CC glycosylation).
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 19; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
DB 6 AAPL 9

RESULT 19
AAY46545
ID AAY46545 standard; Peptide; 9 AA.
XX
AC AAY46545;
XX
DT 01-DEC-1999 (first entry)
XX
DE Immunogenic peptide having a human leukocyte antigen binding motif #1156.
XX
KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW Immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
XX Synthetic.
OS Homo sapiens.
OS
XX WO9445954-A1.
XX
PN 16-SEP-1999.
XX
PD 13-MAR-1998; 98WO-US05039.
XX
PE 13-MAR-1998; 98WO-US05039.
XX
PR 13-MAR-1998; 98WO-US05039.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX
DR WPI; 1999-551214/46.
XX
XX
PT New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases -
XX
XX
PS Claim 1; Page 76; 150pp; English.
XX
CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat
CC or prevent viral infections and cancers in mammals (especially humans)

CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to
CC produce CTLs ex vivo for infusion back into a patient. The
CC polynucleotides encoding the immunogenic peptides are also useful
CC therapeutically and for immunisation as above.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 19; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
DB 5 AAPL 8

RESULT 20
AAY46769
ID AAY46769 standard; Peptide; 9 AA.
XX
AC AAY46769;
XX
DT 01-DEC-1999 (first entry)
XX
DE Immunogenic peptide having a human leukocyte antigen binding motif #1380.
XX
KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW Immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
XX Synthetic.
OS Homo sapiens.
OS
XX WO9445954-A1.
XX
PN 16-SEP-1999.
XX
PD 13-MAR-1998; 98WO-US05039.
XX
PE 13-MAR-1998; 98WO-US05039.
XX
PR 13-MAR-1998; 98WO-US05039.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX
DR WPI; 1999-551214/46.
XX
XX
PT New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases -
XX
XX
PS Claim 1; Page 84; 150pp; English.
XX
CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat
CC or prevent viral infections and cancers in mammals (especially humans)
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in

CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to
CC produce CTLs ex vivo for infusion back into a patient. The
CC polynucleotides encoding the immunogenic peptides are also useful
CC therapeutically and for immunisation as above.

XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 19; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 2 AAPL 5

RESULT 21
AAV46771
ID AAV46771 standard; Peptide: 9 AA.
XX
AC AAV46771;
XX
DT 01-DEC-1999 (first entry)
XX
DE Immunogenic peptide having a human leukocyte antigen binding motif #1382.
XX
KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9945954-A1.
XX
PD 16-SEP-1999.
XX
PF 13-MAR-1998; 98WO-US05039.
XX
PR 13-MAR-1998; 98WO-US05039.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX
DR WPI; 1999-551214/46.
XX
PT New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases -
XX
PS Claim 1; Page 84; 150pp; English.

CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat
CC or prevent viral infections and cancers in mammals (especially humans)
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also

CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to
CC produce CTLs ex vivo for infusion back into a patient. The
CC polynucleotides encoding the immunogenic peptides are also useful
CC therapeutically and for immunisation as above.

XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 19; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 2 AAPL 5

RESULT 22
AAV47521
ID AAV47521 standard; Peptide: 9 AA.
XX
AC AAV47521;
XX
DT 01-DEC-1999 (first entry)
XX
DE Immunogenic peptide having a human leukocyte antigen binding motif #2132.
XX
KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9945954-A1.
XX
PD 16-SEP-1999.
XX
PF 13-MAR-1998; 98WO-US05039.
XX
PR 13-MAR-1998; 98WO-US05039.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX
DR WPI; 1999-551214/46.
XX
PT New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases -
XX
PS Claim 1; Page 113; 150pp; English.

CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat
CC or prevent viral infections and cancers in mammals (especially humans)
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to

CC produce CTLs ex vivo for infusion back into a patient. The
CC polynucleotides encoding the immunogenic peptides are also useful
CC therapeutically and for immunisation as above.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 19; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 6 AAPL 9

RESULT 23
AAV47770
ID AAV47770 standard; Peptide; 9 AA.
XX
AC AAV47770;
XX
DT 01-DEC-1999 (first entry)
XX
DE Immunogenic peptide having a human leukocyte antigen binding motif #2381.
XX
KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW Immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO945954-A1.
XX
PD 16-SEP-1999.
XX
PF 13-MAR-1998; 98WO-US05039.
XX
PR 13-MAR-1998; 98WO-US05039.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
DR WPI; 1999-551214/46.
XX
PT New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases -
XX
PS Claim 1; Page 123; 150pp; English.
XX
CC AAY45390 to AAV48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat
CC or prevent viral infections and cancers in mammals (especially humans)
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to
CC produce CTLs ex vivo for infusion back into a patient. The
CC polynucleotides encoding the immunogenic peptides are also useful

CC therapeutically and for immunisation as above.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 19; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 4 AAPL 7

RESULT 24
AAB48605
ID AAB48605 standard; peptide; 9 AA.
XX
AC AAB48605;
XX
DT 28-FEB-2001 (first entry)
XX
DE C-Abl Tyr kinase activity detection probe motif, SEQ ID NO:4.
XX
KW post-translational modification activity; detection; optical probe;
KW fluorescence; drug screening; sequence specificity determination;
KW protease cleavage site; recognition motif; drug discovery;
KW peptide probe; phosphorylation; tyrosine kinase activity detection;
KW chymotrypsin cleavage.
XX
OS Synthetic.
XX
PN WO20006766-A1.
XX
PD 09-NOV-2000.
XX
PF 05-MAY-2000; 2000WO-US12290.
XX
PR 05-MAY-1999; 99US-0306542.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Pollok BA, Hamman BD, Rodems SM, Makings LR;
DR WPI; 2000-665337/64.
XX
PT Engineered optical probe for measuring a post-translational type
PT modification, comprises a polypeptide with a recognition motif that
PT modulates the polypeptide cleavage rate by a protease, and a
PT fluorescent moiety -
XX
PS Example 1; Page 21; 101pp; English.
XX
CC The invention relates to a engineered optical probe for measuring a
CC post-translational modification activity (e.g., phosphorylation). The
CC probe comprises a non-naturally occurring polypeptide with a recognition
CC motif for the post-translational modification activity, and a protease
CC cleavage site. The probe is attached to a fluorescent moiety.
CC Modification of the polypeptide by the post-translational modification
CC activity results in a modulation of the rate at which a protease cleaves
CC the polypeptide. This is detected by a measurable change in at least one
CC optical property of the optical probe upon cleavage. The invention also
CC includes a recombinant nucleic acid molecule that encodes an optical
CC probe, and vector and host cells comprising this nucleic acid sequence.
CC The invention also relates to a method of determining whether a sample
CC contains a post-translational modification using the probe of the
CC invention; a method of screening a test compound for its ability to
CC modulate a post translational modification activity; a method of
CC determining the sequence specificity of the post-translational activity
CC using a library of probes; systems for spectroscopic measurements
CC utilising a probe of the invention; and therapeutic compositions
CC comprising a modulator of a post-translational activity. The optical
CC probe of the invention is used as an optical sensor of post-translational
CC activities to determine whether a sample contains a post-translational

CC modification, to determine whether a test chemical modulates a post-
CC translational modifying activity, to determine the sequence specificity
CC of a post-translational activity, to identify a therapeutic composition,
CC and in systems for spectroscopic measurements. Compositions containing
CC the probe allow for high throughput and miniaturised screening systems
CC for drug discovery and profiling. Assays using the probe exhibit a large
CC dynamic range, increased sensitivity and allow radiometric readouts for
CC the detection of post-translational activities. The present sequence
CC represents a tyrosine kinase activity detection probe motif which is
CC cleavable by chymotrypsin.
SQ Sequence 9 AA;
QY 1 AAPL 4
Db 6 AAPL 9
Query Match 100.0%; Score 19; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 25
AAU06954
ID AAU06954 standard; Peptide; 9 AA.
XX
AC AAU06954;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human MHC molecule HLA-B7 binding 84P2A9 peptide #22.
XX
KW 84P2A9-related protein; prostate; testis; tissue; cancer; leukaemia;
KW tumour; kidney; brain; bone; skin; ovary; breast; pancreas; colon; lung;
KW cytostatic; gene therapy; antibody therapy; ribozyme; serum; blood;
KW single chain monoclonal antibody; urine.
XX
OS Homo sapiens.
XX
PN WO200155391-A2.
XX
PD 02-AUG-2001.
XX
PF 26-JAN-2001; 2001WO-US02651.
XX
PR 26-JAN-2000; 2000US-0178560.
XX
PA (UROG-) UROGENESYS INC.
XX
PI Jakobovits A, Afar DEH, Challita-eid PM, Levin E, Mitchell SC;
PI Hubert RS;
XX
DR WPI: 2001-502631/55.
XX
PT New 84P2A9 gene and its encoded protein, useful for diagnosing and
PT treating cancer, e.g. leukaemia and cancer of the prostate, testis,
PT kidney, brain or bone, or for eliciting an immune response -
XX
PS Example 12; Page 108; 149pp; English.
XX
CC The polypeptide sequences represent the 84P2A9-related protein and
CC peptide fragments of the protein. 84P2A9 exhibits prostate and testis
CC specific expression in normal adult tissue, but it is also aberrantly
CC expressed in many cancers including leukaemia and tumours of the
CC prostate, testis, kidney, brain, bone, skin, ovary, breast, pancreas,
CC colon and lung. The 84P2A9 polynucleotide, its related protein and
CC peptide fragments and specific PCR primers are therefore useful for
CC diagnosing and treating cancer. A vector comprising a polynucleotide
CC which encodes a single chain monoclonal antibody, that immunospecifically
CC binds to an 84P2A9-related protein, and a ribozyme capable of cleaving a
CC polynucleotide having the 84P2A9 coding sequence, are both useful in the
CC preparation of a composition for treating a patient with a cancer that
CC expresses 84P2A9. The sequences can be used in diagnostic methods to

CC monitor the level of 84P2A9 gene products in serum, blood, urine and
CC tissue and to thereby detect the presence of cancerous cells.
SQ Sequence 9 AA;
QY 1 AAPL 4
Db 6 AAPL 9
Query Match 100.0%; Score 19; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 26
AAU02273
ID AAU02273 standard; Peptide; 9 AA.
XX
AC AAU02273;
XX
DT 29-AUG-2001 (first entry)
XX
DE HLA binding TADG-16 peptide #49.
XX
DE Human; extracellular serine protease; tumour antigen derived gene-16;
KW TADG-16; ovarian carcinoma; breast cancer; lung cancer; colon cancer;
KW prostate cancer; HLA type.
XX
OS Homo sapiens.
XX
PN WO200127257-A1.
XX
PD 19-APR-2001.
XX
PF 13-OCT-2000; 2000WO-US28558.
XX
PR 14-OCT-1999; 99US-0418527.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
PI O'Brien TJ, Underwood LJ, Shigemasa K;
PI WPI: 2001-273769/28.
XX
DR
XX
PT New tumour antigen-derived gene-16 protein, useful for diagnosis and
PT treatment of ovarian, breast, lung, colon and prostate cancer -
XX
PS Example 8; Page 53; 124pp; English.
XX
CC AAU02225-AAU02384 represent TADG-16 peptides which are tested for
CC their binding affinity to the 8 haplotypes HLA A0201, HLA A0205,
CC HLA A1, HLA A24, HLA B7, HLA B8, HLA B2702, and HLA B4403. Tumour
CC antigen derived gene-16 protein, TADG-16 (AAU02223), is a novel human
CC extracellular serine protease. TADG-16 is expressed in normal ovaries
CC and testes and in certain ovarian carcinomas. TADG-16 contains the
CC conserved catalytic triad, His-Asp-Ser, and a signal secretion sequence
CC characteristic of the serine protease family. An antisense
CC oligonucleotide having a complementary sequence to the TADG-16 nucleic
CC acid is useful for treating various cancers, including ovarian, breast,
CC lung, colon and prostate. The TADG-16 nucleic acid, TADG-16 protein and
CC antibodies specific to TADG-16 are useful for the diagnosis of cancer.
CC TADG-16 protein or its fragments are useful for vaccinating an individual
CC against TADG-16.
SQ Sequence 9 AA;
QY 1 AAPL 4
Db 4 AAPL 7
Query Match 100.0%; Score 19; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 27
AAR96502
ID AAR96502 standard; peptide; 10 AA.
XX
AC AAR96502;
XX
DT 07-MAR-1997 (first entry)
XX
DE Hepatitis C virus type 4k peptide.
XX
KW Hepatitis C virus; subtype; polymerase chain reaction; amplification;
KW PCR; primer; probe; antibody; infection.
XX
OS Synthetic.
XX
PN WO9613590-A2.
XX
PD 09-MAY-1996.
XX
PE 23-OCT-1995; 95WO-EP04155.
XX
PR 28-JUN-1995; 95EP-0870076.
PR 21-OCT-1994; 94EP-0870166.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Maertens G, Stuyver L;
XX
DR WPI; 1996-251460/25.
XX
PT Hepatitis C virus poly:nucleic acid unique to unidentified sub:type
PT - used to develop probes and primers for new sub:types and vaccines
PT to prevent and treat infection
XX
PS Claim 5; Page 67; 150pp; English.
XX
CC The peptides AAR96424-R96524 represent novel peptides derived from the
CC novel hepatitis C virus subtypes 1d-f, 2e-i, 2k, 2l, 3g, 4k-m, 7a-c or
CC types 9, 10 or 11 (see AAT27937-T27989). The sequences corresp. to the
CC 5' untranslated region (UR), the Core/EI, NS4 or NS5B regions of the
CC genome. This sequence is from the HCV subtype 4k.
CC The new HCV types were isolated from patients with chronic HCV from the
CC Benelux countries, France, Cameroon and Vietnam, because of their
CC aberrant reactivities. The RNA was extracted, cDNA synthesised and PCR
CC amplified, cloned and genotyped. The 5'UR, Core/EI and NS5B regions were
CC sequenced either directly or partially and used to classify the new
CC viruses into (sub)types based on comparison with known sequences.
CC The nucleotide sequences can be used to synthesise probes and primers
CC for the detection of HCV in a sample. The polypeptides can be used to
CC detect anti-HCV antibodies, for HCV typing or to prevent HCV infections.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 19; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
 ||||
Db 5 AAPL 8

RESULT 28
AAW16482
ID AAW16482 standard; Peptide; 10 AA.
XX
AC AAW16482;
XX
DT 23-JUL-1997 (first entry)
XX
DE Acceptor peptide for N-acetylgalactosaminyltransferase.

XX
KW Acceptor peptide; N-acetylgalactosaminyltransferase; GalNAcT;
KW glycosylation.
XX
OS Synthetic.
XX
PN WO9713783-A1.
XX
PD 17-APR-1997.
XX
PE 09-SEP-1996; 96WO-US14136.
XX
PR 09-OCT-1995; 95US-0005006.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Elhammer AP, Kurosaka A;
XX
DR WPI; 1997-235838/21.
XX
PT Acceptor polypeptide for N-acetylgalactosaminyltransferase
PT useful to develop products for altering glycosylation of proteins
PT and peptide(s)
XX
PS Claim 1; Page 63; 92pp; English.
XX
CC A novel synthetic acceptor peptide (I) (AAW16482) for the enzyme
CC N-acetylgalactosaminyltransferase (GalNAcT) (see also AAW16484) can be
CC used to control glycosylation of a protein or peptide. The
CC acceptor specificity of bovine colostrum GalNAcT (see also AAW16484)
CC was detd. using a semi-quantitative analysis of the amino acids
CC surrounding known glycosylation sites in 16 different proteins.
CC Synthetic acceptor peptides (see also AAW16485-88) were then designed
CC and the transfer of 3H-acetylgalactosamine to the acceptor peptides
CC by GalNAcT was examined. The results showed that peptide (I) had
CC the highest catalytic efficiency of the peptides examined. The
CC glycosylation of a protein may be controlled by incorporating a
CC DNA sequence encoding (I) into the gene for the protein, and
CC exposing the expressed protein to GalNAcT.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 19; DB 18; Length 10;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
 ||||
Db 7 AAPL 10

RESULT 29
AAV46558
ID AAV46558 standard; Peptide; 10 AA.
XX
AC AAV46558;
XX
DT 01-DEC-1999 (first entry)
XX
DE Immunogenic peptide having a human leukocyte antigen binding motif #1169.
XX
KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9945954-A1.
XX
PD 16-SEP-1999.

```
XX 13-MAR-1998; 98WO-US05039.
PF 13-MAR-1998; 98WO-US05039.
PR (EPIM-) EPIMMUNE INC.
PA Settle A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX WPI; 1999-551214/46.
DR
XX New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases -
XX
PS Claim 1; Page 76; 150pp; English.
XX
CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat
CC or prevent viral infections and cancers in mammals (especially humans)
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to
CC produce CTLs ex vivo for infusion back into a patient. The
CC polynucleotides encoding the immunogenic peptides are also useful
CC therapeutically and for immunisation as above.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 19; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 5 AAPL 8

RESULT 30
AAU06981
ID AAU06981 standard; Peptide; 10 AA.
XX
AC AAU06981;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human MHC molecule HLA-B7 binding 84P2A9 peptide #49.
XX
KW 84P2A9-related protein; prostate; testis; tissue; cancer; leukaemia;
KW tumour; kidney; brain; bone; skin; ovary; breast; pancreas; colon; lung;
KW cytostatic; gene therapy; antibody therapy; ribozyme; serum; blood;
KW single chain monoclonal antibody; urine.
XX
OS Homo sapiens.
XX
PN WO200155391-A2.
XX
PD 02-AUG-2001.
XX
PF 26-JAN-2001; 2001WO-US02651.
XX
PR 26-JAN-2000; 2000US-0178560.
```

```
XX (UROG-) UROGENESIS INC.
PA Jakobovits A, Afar DEH, Challita-eid PM, Levin E, Mitchell SC;
XX Hubert RS;
PI WPI; 2001-502631/55.
XX
DR
XX New 84P2A9 gene and its encoded protein, useful for diagnosing and
PT treating cancer, e.g. leukaemia and cancer of the prostate, testis,
PT kidney, brain or bone, or for eliciting an immune response -
XX
PS Example 12; Page 110; 149pp; English.
XX
CC The polypeptide sequences represent the 84P2A9-related protein and
CC peptide fragments of the protein. 84P2A9 exhibits prostate and testis
CC specific expression in normal adult tissue, but it is also aberrantly
CC expressed in many cancers including leukaemia and tumours of the
CC prostate, testis, kidney, brain, bone, skin, ovary, breast, pancreas,
CC colon and lung. The 84P2A9 polynucleotide, its related protein and
CC peptide fragments and specific PCR primers are therefore useful for
CC diagnosing and treating cancer. A vector comprising a polynucleotide
CC which encodes a single chain monoclonal antibody, that immunospecifically
CC binds to an 84P2A9-related protein, and a ribozyme capable of cleaving a
CC polynucleotide having the 84P2A9 coding sequence, are both useful in the
CC preparation of a composition for treating a patient with a cancer that
CC expresses 84P2A9. The sequences can be used in diagnostic methods to
CC monitor the level of 84P2A9 gene products in serum, blood, urine and
CC tissue and to thereby detect the presence of cancerous cells.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 19; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 7 AAPL 10

RESULT 31
AAG94730
ID AAG94730 standard; Peptide; 10 AA.
XX
AC AAG94730;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 924.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04776.
XX
PR 13-DEC-1999; 99GB-0029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
XX A set of peptide ligands consisting of specific complementary peptides
PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
PT drug candidates or pro-drugs -
```

XX
PS Example 4; Page 171; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 19; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
 ||||
Db 5 AAPL 8

RESULT 32
AAR42595
ID AAR42595 standard; peptide; 12 AA.
XX
AC AAR42595;
XX
DT 11-DEC-1994 (first entry)
XX
DE Peptide JBL analogous to D-domain portion of IGF-1.
XX
KW Insulin-like growth factor-1; IGF-1; receptor; antagonist;
KW autophosphorylation inhibitor; C-domain; D-domain; restenosis;
KW cancer; metastasis; asthma; cell proliferation; burns; wounds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 1..12
XX
PN WO9323067-A.
XX
PD 25-NOV-1993.
XX
PF 07-MAY-1993; 93WO-US04329.
XX
PR 08-MAY-1992; 92US-0881524.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Baserga R, Jameson BA;
XX
DR WPI; 1993-386212/48.
XX
PT Compsn. contg. peptide(s) corresp. to C or D domains of
PT insulin-like growth factor - inhibit receptor
PT auto-phosphorylation and used for treating cell proliferation,
PT e.g. cancer, restenosis or asthma
XX
PS Claims 8, 17; Pages 29, 30; 41pp; English.
XX
CC The invention relates to a pharmaceutical composition and method for
CC inhibiting cell proliferation, in which the the active ingredient
CC is a synthetic peptide of less than 25 amino acids comprising at
CC least a portion of the C- or D-domain of human insulin-like growth
CC factor-1. The peptide inhibits IGF-1 induction of auto-
CC phosphorylation by IGF-1 receptor. Diseases which the peptide
CC can be used to treat include restenosis of the coronary arteries
CC after angioplasty, human neoplasia such as cancer of the prostate,
CC tumours in pleural and peritoneal cavities and brain metastases,
CC smooth muscle cell hyperplasia in asthma, burns and wounds, and bone
CC marrow containing highly proliferating cells.
CC The present peptide, designated JBL, includes residues 61-69 in the

CC D-domain of IGF-1 plus additional non-IGF-1 residues including
CC cysteines to form a disulphide bridge to maintain the geometry of the
CC peptide.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 19; DB 14; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
 ||||
Db 3 AAPL 6

RESULT 33
AAW51870
ID AAW51870 standard; peptide; 12 AA.
XX
AC AAW51870;
XX
DT 15-SEP-1998 (first entry)
XX
DE Peptide sequence which binds to IGF-1 receptor.
XX
KW Peptide tail; dicarboxylic acid; marker; chemical coupling; vaccine;
KW chromatography; immunisation; diagnostic test; nerve reconstruction;
KW conformationally constrained; diagnosis; growth factor; IGF-1; bFGF.
XX
OS Synthetic.
XX
PN EP844252-A2.
XX
PD 27-MAY-1998.
XX
PF 14-NOV-1997; 97EP-0870182.
XX
PR 15-NOV-1996; 96US-0030980.
XX
PA (REMA/) REMACLE J.
XX
PI Delforge D, Remacle J;
XX
DR WPI; 1998-274149/25.
XX
PT New cyclic peptides used in biotechnology - comprise di:carboxylic
PT amino acid and tail for coupling to solid supports
XX
PS Disclosure; Page 6; 16pp; English.
XX
CC The invention relates to cyclic peptides which comprise a dicarboxylic
CC amino acid and at least one tail for subsequent weight coupling on an element
CC comprising a solid support, a high molecular weight compound, a marker
CC and/or one or more other similar or different cyclic peptides. The
CC peptides are used in biotechnology applications, particularly in
CC chromatography, immunisation, development of diagnostic tests, vaccines
CC and pharmaceuticals, for Biacor exots, for development of combinatorial
CC libraries of conformationally constrained peptides or for the
CC development of new biomaterials which allow or improve the binding of
CC specific antibodies, receptors, ligands, cells and tissues. The
CC biomaterial may be used for grafting or binding of specific antibodies,
CC receptors, ligands, cells or tissues and may be used for production or
CC identification of new or known drugs, vaccines, hormones, interferons or
CC cytokines and may be used in the screening of specific cells, in
CC particular for the diagnosis of tumour cells, which may adhere
CC differently than normal cells upon the various types of biomaterial. The
CC biomaterial can be used to allow preferential cells and/or tissues
CC adhesion on specific parts of the biomaterial in order to guide the
CC cells and tissue growth along preferential directions, particularly for
CC specific applications such as nerve reconstruction. The present
CC sequence represents a peptide which can fix on the IGF-1 receptor and
CC can advantageously be used as a tailed cyclic peptide.
XX

SQ Sequence 12 AA;

| | | | | |
|-------------------------|---------|--------------------|-----------|------------|
| Query Match | 100.0%; | Score 19; | DB 19; | Length 12; |
| Best Local Similarity | 100.0%; | Pred. No. 3.5e+02; | | |
| Matches 4; Conservative | 0; | Mismatches 0; | Indels 0; | Gaps 0; |

| | | | |
|----|---|------|---|
| QY | 1 | AAPL | 4 |
| | | | |
| Db | 3 | AAPL | 6 |

RESULT 34
AAB92048
ID AAB92048 standard; peptide; 12 AA.

AC AAB92048;

DT 22-JUN-2001 (first entry)

DE Growth factor peptide SEQ ID NO:1224.

KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
 KW blood component; modification; succinimidyl; maleimido group; amino
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

05 Homo sapiens.

OS Synthetic.

PN WO200069900-A2.

PD 23-NOV-2000.

PF 17-MAY-2000; 2000WO-US13576.

PR 17-MAY-1999; 99US-0134406.

PR 15-OCT-1999; 99US-0159783.

PA (CONJ-) CONJUCHEM INC.

PI Bridon DP, Ezrin AM,

DR WPI; 2001-112059/12.

| PT | Modifying and attaching |
|-----|-------------------------|
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |
| 4 | 4 |
| 5 | 5 |
| 6 | 6 |
| 7 | 7 |
| 8 | 8 |
| 9 | 9 |
| 10 | 10 |
| 11 | 11 |
| 12 | 12 |
| 13 | 13 |
| 14 | 14 |
| 15 | 15 |
| 16 | 16 |
| 17 | 17 |
| 18 | 18 |
| 19 | 19 |
| 20 | 20 |
| 21 | 21 |
| 22 | 22 |
| 23 | 23 |
| 24 | 24 |
| 25 | 25 |
| 26 | 26 |
| 27 | 27 |
| 28 | 28 |
| 29 | 29 |
| 30 | 30 |
| 31 | 31 |
| 32 | 32 |
| 33 | 33 |
| 34 | 34 |
| 35 | 35 |
| 36 | 36 |
| 37 | 37 |
| 38 | 38 |
| 39 | 39 |
| 40 | 40 |
| 41 | 41 |
| 42 | 42 |
| 43 | 43 |
| 44 | 44 |
| 45 | 45 |
| 46 | 46 |
| 47 | 47 |
| 48 | 48 |
| 49 | 49 |
| 50 | 50 |
| 51 | 51 |
| 52 | 52 |
| 53 | 53 |
| 54 | 54 |
| 55 | 55 |
| 56 | 56 |
| 57 | 57 |
| 58 | 58 |
| 59 | 59 |
| 60 | 60 |
| 61 | 61 |
| 62 | 62 |
| 63 | 63 |
| 64 | 64 |
| 65 | 65 |
| 66 | 66 |
| 67 | 67 |
| 68 | 68 |
| 69 | 69 |
| 70 | 70 |
| 71 | 71 |
| 72 | 72 |
| 73 | 73 |
| 74 | 74 |
| 75 | 75 |
| 76 | 76 |
| 77 | 77 |
| 78 | 78 |
| 79 | 79 |
| 80 | 80 |
| 81 | 81 |
| 82 | 82 |
| 83 | 83 |
| 84 | 84 |
| 85 | 85 |
| 86 | 86 |
| 87 | 87 |
| 88 | 88 |
| 89 | 89 |
| 90 | 90 |
| 91 | 91 |
| 92 | 92 |
| 93 | 93 |
| 94 | 94 |
| 95 | 95 |
| 96 | 96 |
| 97 | 97 |
| 98 | 98 |
| 99 | 99 |
| 100 | 100 |

PT -
XX
PS Disclosure, Page 595; 733pp; English.

The present invention describes a modified therapeutic peptide (I) comprising a therapeutically active amino acid region (III) and a reactive group (II) (e.g. succinimidyl and maleimido groups) attached to a less therapeutically active amino acid region (IV), which covalently bonds with amino/hydroxyl/thiol groups on blood components to form a peptidase stabilised therapeutic peptide composed of 3-50 amino acids. (I) are useful for modifying therapeutic peptides e.g. hormones, growth factors and neurotransmitters, to protect them from peptidase activity in vivo for the treatment of various disorders. Endogenous therapeutic peptides are not suitable as drug candidates as they require frequent administration due to rapid degradation by peptidases in the body. Modifying and attaching therapeutic peptides to albumin prevents or reduces the action of peptidases to increase length of activity (half life) and specificity as bonding to large molecules decreases intracellular uptake and interference with physiological processes. AAB90829 to AAB92441 represent peptides which can be used in the exemplification of the present invention.

Sequence 12 AA;

| | | | | |
|-----------------------|---------|--------------------|--------|---------------|
| Query Match | 100.0%; | Score 19; | DB 22; | length 12; |
| Best Local Similarity | 100.0%; | Pred. No. 3.5e+02; | | |
| Matches | 4; | Conservative | 0; | Mismatches 0; |
| | | | | Indels 0; |
| | | | | Gaps 0; |

| | | | |
|----|---|------|---|
| Qy | 1 | AAPL | 4 |
| | | | |
| Db | 3 | AAPL | 6 |

RESULT 35
AAB09773
ID AAB09773 standard; peptide; 13 AA

AC AAB09773;

DT 01-SEP-2000 (first entry)

DE Insulin like growth factor related amino acid sequence SEQ ID NO:142.

KW Human; insulin like growth factor; IGF; IGFBP; binding domain;.
KW Insulin like growth factor binding protein; diabetic complication;
KW ischaemic injury; antagonist; modification; gene therapy; cytostatic;
KW vasotropic; antidiabetic; antiParkinsonian; neuroprotective; osteopathic;
KW antiarthritic; vulnerary; tranquilliser; neurologic disease; head trauma;
KW Parkinson's disease; amyotrophic lateral sclerosis; multiple sclerosis;
KW osteoporosis; arthritis.

05 Homo sapiens.

PN WO200023469-A2.

PD 27-APR-2000.

PF 14-OCT-1999; 99WO-US23839.

PR 16-OCT-1998; 98US-0104528.

PA (MUSC-) MUSC FOUND RES DEV

PI Rosenzweig SA, Horney MJ;

DR WPI; 2000-339652/29.

PT New isolated peptide having an insulin-like growth factor domain of an
PT insulin-like growth factor binding protein, useful for treating or
PT preventing cancer or diabetic complications, or for treating ischemic
PT injury -
XX
PS Disclosure; Page 20; 106pp; English.

CC The present invention describes an isolated peptide (A) comprising an
CC insulin like growth factor (IGF) binding domain of an IGF-binding
CC protein (IGFBP) or its modification. (A) binds IGF with at least the same
CC binding affinity as the full length IGFBP. A peptide from the present
CC invention can have cytostatic, vasotropic, antidiabetic,
CC antiParkinsonian, neuroprotective, osteopathic, antiarthritic, vulnery
CC and tranquilliser activities. The peptide is an IGF inhibitor, IGF
CC antagonist and can be used in gene therapy. The peptide and antagonists
CC from the present invention are useful for the treatment or prevention of
CC cancer or diabetic complications, and for treating ischemic injury.
CC Other diseases or injuries that can be treated with the fragment or
CC antagonist include neurologic diseases and injuries, e.g. Parkinson's
CC disease, amyotrophic lateral sclerosis, head trauma or multiple
CC sclerosis, osteoporosis or arthritis. The biotinylated IGF is useful in
CC therapeutic assays for IGFBP and in screening for IGFBP-mimetics (e.g.
CC IGF antagonists). AAB09616 to AAB09773 represent amino acid sequences
CC used in the exemplification of the present invention.

SQ Sequence 13 AA;

| | | | | |
|-------------------------|---------|--------------------|-----------|------------|
| Query Match | 100.0%; | Score 19; | DB 21; | Length 13; |
| Best Local Similarity | 100.0%; | Pred. No. 3.8e+02; | | |
| Matches 4; Conservative | 0; | Mismatches 0; | Indels 0; | Gaps 0 |

QY 1 APR 4
||||

Db 4 AAPL 7

RESULT 36
AAR38723
ID AAR38723 standard; peptide; 14 AA.
XX
AC AAR38723;
XX
DT 09-DEC-1993 (first entry)
XX
DE N-terminal fragment of Megakaryocyte amplifying factor.
XX
KW Megakaryocyte potentiator; Meg-Pot; thrombocytopenia;
KW platelet hypofunction; thrombocyte; interleukin-3; IL-3.
XX
PN WO9313132-A.
XX
PD 08-JUL-1993.
XX
PE 24-DEC-1992; 92WO-JP01689.
XX
PR 27-DEC-1991; 91JP-0361522.
PR 31-MAR-1992; 92JP-0122518.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Hattori K, Oh-Eda M, Yamaguchi N;
XX
DR WPI; 1993-227274/28.
XX
PT Mega-karyocyte amplifier of specified sequence - for treating
PT thrombocytopenia and platelet hypo-function
XX
PS Claim 1; Page 16; 23pp; Japanese.
XX
CC Megakaryocyte potentiator ("Meg-Pot") amplifies megakaryocytes in
CC vitro in the presence of interleukin-3. The protein elutes in the 40-
CC 45% acetonitrile/TFA fraction in reverse phase HPLC, has mol. wt.
CC 32000 (SDS-PAGE) and includes the amino acid sequence AAR38723.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 19; DB 14; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 7 AAPL 10

RESULT 37
AAR53987
ID AAR53987 standard; peptide; 14 AA.
XX
AC AAR53987;
XX
DT 08-DEC-1994 (first entry)
XX
DE Meg-Pot N-terminal peptide.
XX
KW Megakaryocyte potentiator; Meg-Pot; thrombocytopenia;
KW platelet.
XX
OS Homo sapiens.
XX
PN WO9410312-A.
XX
PD 11-MAY-1994.
XX
PF 25-OCT-1993; 93WO-JP01540.
XX

PR 23-OCT-1992; 92JP-0286153.
PR 11-NOV-1992; 92JP-0301387.
PR 09-DEC-1992; 92JP-0329546.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Hattori K, Kojima T, Oh-eda M, Yamaguchi N;
XX
DR WPI; 1994-167467/20.
XX
PT New mega:karyocyte potentiator - for potential treatment of
PT thrombocytopenia
XX
PS Disclosure; Page 49; 74pp; Japanese.
XX
CC DNA encoding Meg-Pot has potential use in treatment of
CC thrombocytopenia and low platelet function.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 19; DB 15; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 7 AAPL 10

RESULT 38
ABB56859
ID ABB56859 standard; Peptide; 14 AA.
XX
AC ABB56859;
XX
DT 05-MAR-2002 (first entry)
XX
DE Human SNP related amino acid sequence SEQ ID NO:1424.
XX
KW Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW autoimmune disease; inflammation; cancer; nervous system disease;
KW infection; polymorphic protein.
XX
OS Homo sapiens.
XX
PN WO200138586-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32311.
XX
PR 24-NOV-1999; 99US-0167383.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-355949/37.
XX
PT Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a
PT pathology, e.g. autoimmune diseases, ascribed to the presence of a
PT sequence polymorphism -
XX
PS Claim 1; Page 664; 674pp; English.
XX
CC ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC to ABB56903 represent human peptides encoded by some of the SNP
CC oligonucleotides. The sequences from the present invention can have
CC immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides

CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterised
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the
CC polymorphic protein within appropriate physiological samples).

SQ Sequence 14 AA;

Query Match 100.0%; Score 19; DB 22; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
| | | |
Db 1 AAPL 4

RESULT 39

AAB67038 ID AAB67038 standard; Peptide; 14 AA.

AC AAB67038;

DT 09-APR-2001 (first entry)

DE Human insulin-like growth factor agonist peptide SEQ ID NO: 38.

XX Human; insulin-like growth factor; IGF; agonist; hyperglycaemic disorder;
KW obesity; neurological disorder; cardiac disorder; renal disorder;
KW immunological disorder; anabolic disorder.

OS Homo sapiens.

PN WO200078801-A2.

PD 28-DEC-2000.

PF 19-JUN-2000; 2000WO-US17023.

PR 22-JUN-1999; 99US-0337227.

PA (GETH) GENENTECH INC.

PI Chen YM, Cochran AG, Lowman HB, Skelton NJ;

DR WPI; 2001-112312/12.

XX New peptide for increasing serum and tissue levels of biological active

PT insulin growth factor -

PS Example 1; Page 54; 75pp; English.

CC The present invention provides the sequences of a number of peptides
CC which act as human insulin growth factor (IGF) agonists. These can be
CC used in the treatment of hyperglycaemic, obesity-related, neurological,
CC cardiac, renal, immunological and anabolic disorders.

SQ Sequence 14 AA;

Query Match 100.0%; Score 19; DB 22; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
| | | |
Db 3 AAPL 6

RESULT 40

ABP46386 ID ABP46386 standard; peptide; 14 AA.

AC ABP46386;

DT 19-AUG-2002 (first entry)

DE Human Blys binding scFv VH CDR3 SEQ ID 2397.

XX Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.

OS Homo sapiens.

PN WO200202641-A1.

PD 10-JAN-2002.

PF 15-JUN-2001; 2001WO-US19110.

PR 16-JUN-2000; 2000US-212210P.

PR 17-OCT-2000; 2000US-240816P.

PR 16-MAR-2001; 2001US-276248P.

PR 21-MAR-2001; 2001US-277379P.

PR 25-MAY-2001; 2001US-293499P.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.

PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

DR WPI; 2002-114799/15.

XX Antibodies against B lymphocyte stimulating polypeptides, useful for

PT the diagnosis and treatment of cancers and immune disorders -

PS Claim 2; Page 2983; 3148pp; English.

XX This invention describes novel antibodies that immunospecifically bind to
CC B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytosstatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antirheumatic and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of Blys. The antibodies bind to Blys
CC and so may be used to detect and quantitate the presence of Blys in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of Blys. They may also be
CC administered to treat diseases associated with aberrant Blys expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC acquired immunodeficiency syndrome (AIDS)). ABP4390-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method
CC of the invention.

SQ Sequence 14 AA;

Query Match 100.0%; Score 19; DB 23; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
| | | |
Db 9 AAPL 12

Job time : 27 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:26:14 ; Search time 9.33333 Seconds
(without alignments)
12.610 Million cell updates/sec

Title: US-10-033-526-3
Perfect score: 19
Sequence: 1 AAPL 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Query length | DB | ID | Description |
|------------|-------|-------|--------------|----|--------------------|--------------------|
| 1 | 19 | 100.0 | 4 | 1 | US-08-004-643C-4 | Sequence 4, Appli |
| 2 | 19 | 100.0 | 4 | 1 | US-08-345-820B-1 | Sequence 1, Appli |
| 3 | 19 | 100.0 | 4 | 1 | US-08-544-143A-6 | Sequence 6, Appli |
| 4 | 19 | 100.0 | 4 | 1 | US-08-224-868-9 | Sequence 9, Appli |
| 5 | 19 | 100.0 | 4 | 1 | US-08-777-208-6 | Sequence 6, Appli |
| 6 | 19 | 100.0 | 4 | 1 | US-08-693-653-1 | Sequence 1, Appli |
| 7 | 19 | 100.0 | 4 | 2 | US-08-698-575E-1 | Sequence 1, Appli |
| 8 | 19 | 100.0 | 4 | 2 | US-08-907-840A-3 | Sequence 3, Appli |
| 9 | 19 | 100.0 | 4 | 2 | US-08-765-165-1 | Sequence 1, Appli |
| 10 | 19 | 100.0 | 4 | 3 | US-09-069-823-1 | Sequence 10, Appli |
| 11 | 19 | 100.0 | 4 | 4 | US-09-578-303-10 | Sequence 3, Appli |
| 12 | 19 | 100.0 | 6 | 4 | US-08-988-842-3 | Sequence 11, Appli |
| 13 | 19 | 100.0 | 8 | 1 | US-08-057-184-11 | Sequence 12, Appli |
| 14 | 19 | 100.0 | 8 | 1 | US-08-057-184-11 | Sequence 143, App |
| 15 | 19 | 100.0 | 9 | 2 | US-08-340-283-143 | Sequence 29, Appli |
| 16 | 19 | 100.0 | 9 | 4 | US-08-772-282-29 | Sequence 4, Appli |
| 17 | 19 | 100.0 | 9 | 4 | US-09-306-542A-4 | Sequence 51, Appli |
| 18 | 19 | 100.0 | 10 | 3 | US-08-836-075A-185 | Sequence 185, App |
| 19 | 19 | 100.0 | 10 | 4 | US-09-421-208-51 | Sequence 51, Appli |
| 20 | 19 | 100.0 | 10 | 4 | US-08-835-231-2 | Sequence 2, Appli |
| 21 | 19 | 100.0 | 11 | 2 | US-08-967-508-17 | Sequence 17, Appli |
| 22 | 19 | 100.0 | 11 | 2 | US-08-967-506-17 | Sequence 2, Appli |
| 23 | 19 | 100.0 | 11 | 4 | US-09-108-661-2 | Sequence 17, Appli |
| 24 | 19 | 100.0 | 11 | 4 | PCT-US94-02552-17 | Sequence 3, Appli |
| 25 | 19 | 100.0 | 12 | 1 | US-08-219-878A-3 | Sequence 3, Appli |
| 26 | 19 | 100.0 | 12 | 1 | PCT-US93-04329-3 | Sequence 3, Appli |
| 27 | 19 | 100.0 | 12 | 5 | | |

| | | | | | | |
|----|----|-------|----|---|-------------------|--------------------|
| 28 | 19 | 100.0 | 14 | 1 | US-08-256-133-1 | Sequence 1, Appli |
| 29 | 19 | 100.0 | 14 | 1 | US-08-426-819A-4 | Sequence 4, Appli |
| 30 | 19 | 100.0 | 14 | 4 | US-09-337-227C-38 | Sequence 38, Appli |
| 31 | 19 | 100.0 | 16 | 1 | US-08-256-133-3 | Sequence 3, Appli |
| 32 | 19 | 100.0 | 16 | 1 | US-08-256-133-4 | Sequence 5, Appli |
| 33 | 19 | 100.0 | 16 | 1 | US-08-256-133-5 | Sequence 1, Appli |
| 34 | 19 | 100.0 | 16 | 1 | US-08-426-819A-1 | Sequence 1, Appli |
| 35 | 19 | 100.0 | 16 | 1 | US-08-426-819A-2 | Sequence 3, Appli |
| 36 | 19 | 100.0 | 16 | 1 | US-08-426-819A-3 | Sequence 4, Appli |
| 37 | 19 | 100.0 | 19 | 1 | US-08-735-963-4 | Sequence 4, Appli |
| 38 | 19 | 100.0 | 19 | 2 | US-09-105-057-4 | Sequence 4, Appli |
| 39 | 19 | 100.0 | 19 | 4 | US-09-304-214-4 | Sequence 26, Appli |
| 40 | 19 | 100.0 | 20 | 2 | US-08-124-981A-26 | Sequence 110, App |
| 41 | 19 | 100.0 | 20 | 3 | US-08-840-316-110 | Sequence 111, App |
| 42 | 19 | 100.0 | 20 | 3 | US-08-840-316-111 | Sequence 110, App |
| 43 | 19 | 100.0 | 20 | 4 | US-09-402-776-110 | Sequence 111, App |
| 44 | 19 | 100.0 | 20 | 4 | US-09-402-776-111 | Sequence 1, Appli |
| 45 | 19 | 100.0 | 25 | 2 | US-08-745-881-1 | |

ALIGNMENTS

RESULT 1
US-08-004-643C-4
; Sequence 4, Application US/08004643C
; Patent No. 5480779
; GENERAL INFORMATION:
; APPLICANT: Gunter Fischer & Gerhard K llerz
; TITLE OF INVENTION: Cyclosporine Assay
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Gunter Fischer et al. c/o G. P. Katona
; STREET: 230 Park Avenue, Room 2200
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10169
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: conf. to Patentin Release #1.0, Ver.#1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/004,643C
; FILING DATE: 12 January 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 703,590
; FILING DATE: 20 May 1991
; APPLICATION NUMBER: 398,092
; FILING DATE: 24 August 1989
; APPLICATION NUMBER: DD WP 601 F/319 577W
; FILING DATE: 07 September 1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Katona, Gabriel P.
; REGISTRATION NUMBER: 20,829
; REFERENCE/DOCKET NUMBER: 691-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-3377
; TELEFAX: (212)986-6126
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-004-643C-4
Query Match 100.0%; Score 19; DB 1; Length 4;
Best local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 2

US-08-345-820B-1
; Sequence 1, Application US/08345820B
; Patent No. 5618792
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: SUBSTITUTED HETEROCYCLIC COMPOUNDS USEFUL AS
; TITLE OF INVENTION: INHIBITORS OF (SERINE PROTEASES) HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/345,820B
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-345-820B-1

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 3

US-08-544-143A-6
; Sequence 6, Application US/08544143A
; Patent No. 5646028
; GENERAL INFORMATION:
; APPLICANT: Leigh, Scott D.
; TITLE OF INVENTION: NOVEL ALKALINE PROTEASE AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Majestic, Parsons, Siebert & Hsue
; STREET: Four Embarcadero Center, Suite 1100
; CITY: San Francisco
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94111-4121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/544,143A
; FILING DATE: 17-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: de Runtz, K. Alison
; REGISTRATION NUMBER: 37,119
; REFERENCE/DOCKET NUMBER: 0409.054US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-362-5556
; TELEFAX: 415-362-5418
; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-544-143A-6

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 4

US-08-224-868-9
; Sequence 9, Application US/08224868
; Patent No. 5698448
; GENERAL INFORMATION:
; APPLICANT: Soidin, Steven J.
; TITLE OF INVENTION: IMMUNOSUPPRESSIVE DRUG BINDING PROTEINS
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/224,868
; FILING DATE: 08-APR-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/200,404
; FILING DATE: 23-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/782,761
; FILING DATE: 22-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/487,115
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/279,176
; FILING DATE: 02-DEC-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/841,792
; FILING DATE: 26-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/521,074
; FILING DATE: 09-MAY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 64688/125/CHRE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-224-868-9

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 5

US-08-777-208-6
; Sequence 6, Application US/08777208
; Patent No. 5763576
; GENERAL INFORMATION:
; APPLICANT: Powers, James C.
; TITLE OF INVENTION: Tetrapeptide Alpha-Ketamides
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Deveau, Colton & Marquis
; STREET: Two Midtown Plaza, Suite 1400
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Microsoft-Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/777,208
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/539944
; FILING DATE: 06-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Colton, Laurence P.
; REGISTRATION NUMBER: 33371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 875-3555
; TELEFAX: (404) 875-8505
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: no
US-08-777-208-6

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 6

US-08-693-653-1
; Sequence 1, Application US/08693653
; Patent No. 5780439
; GENERAL INFORMATION:
; APPLICANT: Mendy, Francois
; APPLICANT: Kahn, Jean-Maurice
; APPLICANT: Roger, Loic

; TITLE OF INVENTION: Improvements in or relating to organic
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sandoz Corporation
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: NJ07936

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,653
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/353,652
; FILING DATE:
; APPLICATION NUMBER: US 07/960,143
; FILING DATE: 13-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Battle, Carl
; REGISTRATION NUMBER: 30,731
; REFERENCE/DOCKET NUMBER: 510-5747
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 503-8532
; TELEFAX: (201) 503-8807
; TELEX: 240867

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Synthetic
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /label= 1a
; OTHER INFORMATION: /note= "glutaryl derivative"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /label= 1b
; OTHER INFORMATION: /note= "-p-nitro-anilide derivative"
US-08-693-653-1

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 7

US-08-698-575E-1
; Sequence 1, Application US/08698575E
; Patent No. 5874585
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: SUBSTITUTED HETEROCYCLIC COMPOUNDS USEFUL AS
; INHIBITORS OF (SERINE PROTEASES) HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 4

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/698,575E
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/345,820
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-698-575E-1

Query Match 100.0%; Score 19; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 8
US-08-907-840A-3
; Sequence 3, Application US/08907840A
; Patent No. 5952307
; GENERAL INFORMATION:
; APPLICANT: Powers, James C.
; TITLE OF INVENTION: Basic Alpha-Aminoalkylphosphonate
; TITLE OF INVENTION: Derivatives
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Deveau, Colton & Marquis
; STREET: Two Midtown Plaza, Suite 1400
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/907,840A
; FILING DATE: 14 AUG-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/184286
; FILING DATE: 21 JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Colton, Laurence P.
; REGISTRATION NUMBER: 33371
; REFERENCE/DOCKET NUMBER: 10733-175CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 875-3555
; TELEFAX: (404) 875-8505
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: no
; US-08-907-840A-3

Query Match 100.0%; Score 19; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 9
US-08-765-165-1
; Sequence 1, Application US/08765165
; Patent No. 5985273
; GENERAL INFORMATION:
; APPLICANT: Reed, Benjamin J.
; APPLICANT: Sandeman, Richard M.
; APPLICANT: Chandler, David S.
; TITLE OF INVENTION: BIOLOGICAL CONTROL OF INSECTS
; FILE REFERENCE: 08820/002001
; CURRENT APPLICATION NUMBER: US/08/765,165
; CURRENT FILING DATE: 1997-06-27
; EARLIER APPLICATION NUMBER: PCT/AU95/00347
; EARLIER FILING DATE: 1995-06-17
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic substrates
; US-08-765-165-1

Query Match 100.0%; Score 19; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 10
US-09-069-823-1
; Sequence 1, Application US/09069823
; Patent No. 6037325
; GENERAL INFORMATION:
; APPLICANT: Gyorkos, Albert C.
; APPLICANT: Spruce, Lyle W.
; TITLE OF INVENTION: SUBSTITUTED HETEROCYCLIC COMPOUNDS
; TITLE OF INVENTION: USEFUL AS INHIBITORS OF (SERINE PROTEASES) HUMAN NEUTROPHIL
; FILE REFERENCE: 20774.240087
; CURRENT APPLICATION NUMBER: US/09/069,823
; CURRENT FILING DATE: 1998-04-30
; EARLIER APPLICATION NUMBER: 08/345,820
; EARLIER FILING DATE: 1994-11-21
; EARLIER APPLICATION NUMBER: 08/698,575
; EARLIER FILING DATE: 1996-08-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Tetrapeptide
; US-09-069-823-1

Query Match 100.0%; Score 19; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 11

US-09-578-303-10
; Sequence 10, Application US/09578303
; Patent No. 6399759
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Whitworth, S. Troy
; APPLICANT: Blum, Murray S.
; TITLE OF INVENTION: Ant Proteases and Methods of Inhibition
; FILE REFERENCE: 235.00150101
; CURRENT APPLICATION NUMBER: US/09/578,303
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/136,331
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: para-nitroanilide substrate
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: succinyl end cap
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: (4)
; OTHER INFORMATION: para-nitroanilide end cap
US-09-578-303-10

Query Match 100.0%; Score 19; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 12
US-08-988-842-3
; Sequence 3, Application US/08988842
; Patent No. 6462173
; GENERAL INFORMATION:
; APPLICANT: Lu, Jun Ping
; APPLICANT: Cantley, Lewis C.
; APPLICANT: Yaffe, Michael
; APPLICANT: Fischer, Gunter
; TITLE OF INVENTION: INHIBITORS OF PHOSPHOSERINE AND
; TITLE OF INVENTION: PHOSPHOTHREONINE-PROLINE-SPECIFIC ISOMERASES
; FILE REFERENCE: BIDMC97-02PA
; CURRENT APPLICATION NUMBER: US/08/988,842
; CURRENT FILING DATE: 1997-12-11
; EARLIER APPLICATION NUMBER: 60/058,164
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 6
; TYPE: PRT
; ORGANISM: synthetic nucleotide
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: (5)...(5)
US-08-988-842-3

Query Match 100.0%; Score 19; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 13
US-08-057-184-11
; Sequence 11, Application US/08057184
; Patent No. 538588
; GENERAL INFORMATION:
; APPLICANT: Goodenow, Robert S
; APPLICANT: Olsson, Lennart
; TITLE OF INVENTION: Class I MHC Modulation of Surface
; TITLE OF INVENTION: Receptor Activity
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bertram I. Rowland
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/057,184
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/649,471
; FILING DATE: 01-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-55115-2/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-057-184-11

Query Match 100.0%; Score 19; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 5 AAPL 8

RESULT 14
US-08-057-184-12
; Sequence 12, Application US/08057184
; Patent No. 538588
; GENERAL INFORMATION:
; APPLICANT: Goodenow, Robert S
; APPLICANT: Olsson, Lennart
; TITLE OF INVENTION: Class I MHC Modulation of Surface
; TITLE OF INVENTION: Receptor Activity
; NUMBER OF SEQUENCES: 17
US-08-057-184-12

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Bertram I. Rowland
;; STREET: 4 Embarcadero Center, Suite 3400
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94111
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/057,184
;; FILING DATE:
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/649,471
;; FILING DATE: 01-FEB-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Rowland, Bertram I
;; REGISTRATION NUMBER: 20,015
;; REFERENCE/DOCKET NUMBER: A-55115-2/BIR
;; TELEPHONE: (415) 781-1989
;; TELEFAX: (415) 398-3249
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 8 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;;
US-08-057-184-12

Query Match 100.0%; Score 19; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 5 AAPL 8

RESULT 15
US-08-340-283-143
;; Sequence 143, Application US/08340283
;; Patent No. 5861318
;;
;; GENERAL INFORMATION:
;; APPLICANT: Elhammer, Ake P.
;; TITLE OF INVENTION: A SCINTILLATION PROXIMITY ASSAY FOR
;; TITLE OF INVENTION: N-ACETYLGLACTOSAMINYLTRANSFERASE ACTIVITY
;; NUMBER OF SEQUENCES: 205
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pharmacia and Upjohn, Inc., Intellect. Prop. Law
;; ADDRESSEE: (1920-32-1)
;; STREET: 301 Henrietta Street
;; CITY: Kalamazoo
;; STATE: Michigan
;; COUNTRY: U.S.A.
;; ZIP: 49001
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/340,283
;; FILING DATE:
;; CLASSIFICATION: 436
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Wootton, Thomas A.
;; REGISTRATION NUMBER: 35,004

;; REFERENCE/DOCKET NUMBER: 4828
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (616) 385-7914
;; TELEFAX: (616) 385-6897
;; TELEX: 224401
;;
;; INFORMATION FOR SEQ ID NO: 143:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;;
US-08-340-283-143

Query Match 100.0%; Score 19; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 6 AAPL 9

RESULT 16
US-08-772-282-29
;; Sequence 29, Application US/08772282
;; Patent No. 6270777
;;
;; GENERAL INFORMATION:
;; APPLICANT: SOKOL, Pamela A.
;; APPLICANT: KOOL, Cora D.
;; TITLE OF INVENTION: CONSERVED METALLOPROTEASE EPTIOPES
;; NUMBER OF SEQUENCES: 29
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
;; STREET: P.O. Box 1404
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: USA
;; ZIP: 22313-1404
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/772,282
;; FILING DATE: 20-DEC-1996
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: STANEK REA, Teresa
;; REGISTRATION NUMBER: 30,427
;; REFERENCE/DOCKET NUMBER: 024916-005
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-2021
;; INFORMATION FOR SEQ ID NO: 29:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;;
US-08-772-282-29

Query Match 100.0%; Score 19; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 1111

```
Db      4 AAPL 7

RESULT 17
; US-09-306-542A-4
; Sequence 4, Application US/09306542A
; Patent No. 6410255
; GENERAL INFORMATION:
; APPLICANT: AURORA BIOSCIENCES CORPORATION
; APPLICANT: POLLOK, Brian
; APPLICANT: HAMMAN, Brian
; APPLICANT: RODAMS, Steven
; APPLICANT: MAKINGS, Lewis
; TITLE OF INVENTION: OPTICAL PROBES AND ASSAYS
; FILE REFERENCE: AURO1300
; CURRENT APPLICATION NUMBER: US/09/306,542A
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Optical probe specific motif
; US-09-306-542A-4

Query Match      100.0%; Score 19; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPL 4
        1111
Db      6 AAPL 9

RESULT 18
; US-09-105-678A-51
; Sequence 51, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Corlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 51:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-51

Query Match      100.0%; Score 19; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPL 4
        1111
Db      5 AAPL 8

RESULT 19
; US-08-836-075A-185
; Sequence 185, Application US/08836075A
; Patent No. 6180768
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: STUYVER, LIEVEN
; TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
; TITLE OF INVENTION: AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
; TITLE OF INVENTION: AGENTS
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word 6.0 / ASCII text output
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,075A
; FILING DATE: 21 Apr 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04155
; FILING DATE: 23 Oct 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94870166.9
; FILING DATE: 21 Oct 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95870076.7
; FILING DATE: 28 Jun 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KAMMERER, PATRICIA A.
; REGISTRATION NUMBER: 29,775
; REFERENCE/DOCKET NUMBER: INNS:004
; INFORMATION FOR SEQ ID NO: 185:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-836-075A-185

Query Match      100.0%; Score 19; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
RESULT 20
US-09-421-208-51
; Sequence 51, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-09-421-208-51
Query Match      100.0%; Score 19; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APL 4
      1111
Db      4 APL 7

RESULT 21
US-08-835-231-2
; Sequence 2, Application US/08835231
; Patent No. 5861284
; GENERAL INFORMATION:
; APPLICANT: NISHIMURA, Osamu
; APPLICANT: KURIYAMA, Masato
; APPLICANT: Koyama, No. 5861284uyuki
; APPLICANT: FUKUDA, Tsunehiko
; TITLE OF INVENTION: METHOD FOR PRODUCING A BIOLOGICALLY
; TITLE OF INVENTION: ACTIVE RECOMBINANT CYSTEINE-FREE
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MA
```

```

; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,231
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,709
; FILING DATE: 07-DEC-1994
; APPLICATION NUMBER: 07/838,857
; FILING DATE: 18-FEB-1992
; APPLICATION NUMBER: JP 024841
; FILING DATE: 19-FEB-1991
; APPLICATION NUMBER: JP 0271438
; FILING DATE: 18-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DAVID, RESNICK S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 41614-FWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX: 200291 STRE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 4
; OTHER INFORMATION: /note = "Xaa = D-Alanine"
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
;
US-08-835-231-2
Query Match      100.0%; Score 19; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APL 4
      1111
Db      5 APL 8

RESULT 22
US-08-967-508-17
; Sequence 17, Application US/08967508
; Patent No. 5910570
; GENERAL INFORMATION:
; APPLICANT: The Upjohn Company
; APPLICANT: FOR U.S. PURPOSES ONLY: Elhammer, Ake P. and Homa, Fred L.
; TITLE OF INVENTION: A Cloned DNA Encoding a UDP-GalNAc:
; TITLE OF INVENTION: Polypeptide, N-Acetylgalactosaminyltransferase
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pharmacia & Upjohn Company, Intellectual
; ADDRESSEE: Property Legal Services
; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49001
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,508
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/602,830
FILING DATE: 13 No. 5910570ember 1995
ATTORNEY/AGENT INFORMATION:
NAME: Darnley Jr., James D.
REGISTRATION NUMBER: 33,673
REFERENCE/DOCKET NUMBER: 4755.P CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 616-833-2210
TELEFAX: 616-833-8897
TELEX: 224401
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-967-508-17

Query Match 100.0%; Score 19; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 7 AAPL 10

RESULT 23
US-08-967-506-17
Sequence 17, Application US/08967506
Patent No. 6096512
GENERAL INFORMATION:
APPLICANT: The Upjohn Company
APPLICANT: FOR U.S. PURPOSES ONLY: Elhammer, Ake P. and Homa, Fred L.
TITLE OF INVENTION: A Cloned DNA Encoding a UDP-GalNAc:
TITLE OF INVENTION: Polypeptide, N-Acetylgalactosaminyltransferase
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pharmacia & Upjohn Company, Intellectual
ADDRESSEE: Property Legal Services
STREET: 301 Henrietta Street
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,506
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/602,830
FILING DATE: 13 No. 6096512ember 1995
ATTORNEY/AGENT INFORMATION:
NAME: Darnley Jr., James D.
REGISTRATION NUMBER: 33,673
REFERENCE/DOCKET NUMBER: 4755.P CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 616-833-2210

TELEFAX: 616-833-8897
TELEX: 224401
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-967-506-17

Query Match 100.0%; Score 19; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 7 AAPL 10

RESULT 24
US-09-108-661-2
Sequence 2, Application US/09108661
Patent No. 6287806
GENERAL INFORMATION:
APPLICANT: NISHIMURA, Osamu
APPLICANT: KURIYAMA, Masato
APPLICANT: KUYAMA, No. 6287806uyuki
APPLICANT: FUKUDA, Tsunehiko
TITLE OF INVENTION: METHOD FOR PRODUCING A BIOLOGICALLY
TITLE OF INVENTION: ACTIVE RECOMBINANT CYSTEINE-FREE
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 WATER STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/108,661
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,709
FILING DATE: 07-DEC-1994
APPLICATION NUMBER: 07/838,857
FILING DATE: 18-FEB-1992
APPLICATION NUMBER: JP 024841
FILING DATE: 19-FEB-1991
APPLICATION NUMBER: JP 0271438
FILING DATE: 18-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: DAVID, RESNICK S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 41614-FWC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX: 200291 STRE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO

; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 4
; OTHER INFORMATION: /note = "Xaa = D-Alanine"
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-09-108-661-2

Query Match 100.0%; Score 19; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 5 AAPL 8

RESULT 25

PCT-US94-02552-17
; Sequence 17, Application PC/TUS9402552

; GENERAL INFORMATION:

; APPLICANT: Elhammer, Ake P.

; APPLICANT: Home, Fred L.

; TITLE OF INVENTION: A Cloned DNA Encoding a UDP-GALNAc:

; TITLE OF INVENTION: Polypeptide, N-Acetylgalactosaminyltransferase

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: The Upjohn Company, Corp. Intellectual

; ADDRESSEE: Property Law

; STREET: 301 Henrietta Street

; CITY: Kalamazoo

; STATE: Michigan

; COUNTRY: USA

; ZIP: 49001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/02552

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Darnley Jr., James D.

; REGISTRATION NUMBER: 33,673

; REFERENCE/DOCKET NUMBER: 4755.P CP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 616-385-5210

; TELEFAX: 616-385-6897

; TELEX: 224401

; INFORMATION FOR SEQ ID NO: 17:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 11 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

PCT-US94-02552-17

Query Match 100.0%; Score 19; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 7 AAPL 10

RESULT 26
US-08-219-878A-3
; Sequence 3, Application US/08219878A

; Patent No. 5473054
; GENERAL INFORMATION:
; APPLICANT: Bradford A. Jameson and Renato Baserga
; TITLE OF INVENTION: IGF-1 Analogs
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn

; ADDRESSEE: Kurtz Mackiewicz & NO. 5473054ris

; STREET: One Liberty Place - 46th Floor

; CITY: Philadelphia

; STATE: PA

; COUNTRY: USA

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/219,878A

; FILING DATE: 30-MAR-1994

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/881,524

; FILING DATE: 08-MAY-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Mark DeLuca

; REGISTRATION NUMBER: 33,229

; REFERENCE/DOCKET NUMBER: TJU-1240

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-3100

; TELEFAX: (215) 568-3439

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 12

; TYPE: amino acid

; TOPOLOGY: linear

US-08-219-878A-3

Query Match 100.0%; Score 19; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 3 AAPL 6

RESULT 27

PCT-US93-04329-3
; Sequence 3, Application PC/TUS9304329

; GENERAL INFORMATION:

; APPLICANT: Bradford A. Jameson and Renato Baserga

; TITLE OF INVENTION: IGF-1 Analogs

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn

; ADDRESSEE: Kurtz Mackiewicz & Norris

; STREET: One Liberty Place - 46th Floor

; CITY: Philadelphia

; STATE: PA

; COUNTRY: USA

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: WORDPERFECT 5.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/04329

; FILING DATE: 19930507

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/881,524
FILING DATE: 08-MAY-92,
ATTORNEY/AGENT INFORMATION:
NAME: Mark Deluca
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-0649
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 12
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US93-04329-3

Query Match 100.0%; Score 19; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 3 AAPL 6

RESULT 28
US-08-256-133-1
Sequence 1, Application US/08256133
Patent No. 5498698
GENERAL INFORMATION:
APPLICANT: Yamaguchi, No. 5498698om1
APPLICANT: Oh-Eda, Masayoshi
APPLICANT: Hattori, Kunihiro
TITLE OF INVENTION: No. 5498698e1 Megakaryocyte Potentiator
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,133
FILING DATE: 27-JUN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 230-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-256-133-1

Query Match 100.0%; Score 19; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 7 AAPL 10

RESULT 29
US-08-426-819A-4
Sequence 4, Application US/08426819A
Patent No. 5723318
GENERAL INFORMATION:
APPLICANT: Yamaguchi, No. 5723318om1
APPLICANT: Kojima, Tetsuo
APPLICANT: Oh-Eda, Masayoshi
APPLICANT: Hattori, Kunihiro
TITLE OF INVENTION: Genes Coding for Megakaryocyte
TITLE OF INVENTION: Potentiator
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,819A
FILING DATE: 21-APR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 230-107P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
FRAGMENT TYPE: N-terminal
IMMEDIATE SOURCE:
CLONE: HPCY5
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..14
OTHER INFORMATION: /label= peptide
OTHER INFORMATION: /note= "consensus sequence from Edman runs of amino termin
US-08-426-819A-4

Query Match 100.0%; Score 19; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 7 AAPL 10

RESULT 30

US-09-337-227C-38
: Sequence 38, Application US/09337227C
: Patent No. 6420518
: GENERAL INFORMATION:
: APPLICANT: Chen, Yvonne May-Yee
: APPLICANT: Clark, Ross G.
: APPLICANT: Cochran, Andrea G.
: APPLICANT: Lowman, Henry B.
: APPLICANT: Robinson, Iain C.A.F.
: APPLICANT: Skelton, Nicholas J.
: TITLE OF INVENTION: INSULIN-LIKE GROWTH FACTOR AGONIST MOLECULES
: FILE REFERENCE: P1071P2.rev
: CURRENT APPLICATION NUMBER: US/09/337, 227C
: CURRENT FILING DATE: 1999-06-22
: PRIOR APPLICATION NUMBER: US 09/052, 888
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: US 08/825, 852
: PRIOR FILING DATE: 1997-04-04
: NUMBER OF SEQ ID NOS: 51
: SEQ ID NO 38
: LENGTH: 14
: TYPE: PRT
: ORGANISM: Artificial sequence
: FEATURE:
: OTHER INFORMATION: Sequence is synthesized
: Patent No. 6420518
US-09-337-227C-38

Query Match 100.0%; Score 19; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 3 AAPL 6

RESULT 31
US-08-256-133-3
: Sequence 3, Application US/08256133
: Patent No. 5498698
: GENERAL INFORMATION:
: APPLICANT: Yamaguchi, No. 5498698omi
: APPLICANT: Oh-Eda, Masayoshi
: APPLICANT: Hattori, Kunihiko
: TITLE OF INVENTION: No. 5498698el Megakaryocyte Potentiator
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Birch, Stewart, Kolasch & Birch
: STREET: 8110 Gatehouse Road, Suite 500 East
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22042
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/256, 133
: FILING DATE: 27-JUN-1994
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Murphy Jr., Gerald M.
: REGISTRATION NUMBER: 28,977
: REFERENCE/DOCKET NUMBER: 230-104P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 205-8000
: TELEFAX: (703) 205-8050
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:

: LENGTH: 16 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
US-08-256-133-3

Query Match 100.0%; Score 19; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 7 AAPL 10

RESULT 32
US-08-256-133-4
: Sequence 4, Application US/08256133
: Patent No. 5498698
: GENERAL INFORMATION:
: APPLICANT: Yamaguchi, No. 5498698omi
: APPLICANT: Oh-Eda, Masayoshi
: APPLICANT: Hattori, Kunihiko
: TITLE OF INVENTION: No. 5498698el Megakaryocyte Potentiator
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Birch, Stewart, Kolasch & Birch
: STREET: 8110 Gatehouse Road, Suite 500 East
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22042
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/256, 133
: FILING DATE: 27-JUN-1994
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Murphy Jr., Gerald M.
: REGISTRATION NUMBER: 28,977
: REFERENCE/DOCKET NUMBER: 230-104P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 205-8000
: TELEFAX: (703) 205-8050
: TELEX: 248345
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 16 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
US-08-256-133-4

Query Match 100.0%; Score 19; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 8 AAPL 11

RESULT 33
US-08-256-133-5
; Sequence 5, Application US/08256133
; Patent No. 5498698
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, No. 5498698omi
; APPLICANT: Oh-Eda, Masayoshi
; APPLICANT: Hattori, Kunihiko
; TITLE OF INVENTION: No. 5498698el Megakaryocyte Potentiator
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,133
; FILING DATE: 27-JUN-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 230-104P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-256-133-5

Query Match 100.0%; Score 19; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 9 AAPL 12

RESULT 34
US-08-426-819A-1
; Sequence 1, Application US/08426819A
; Patent No. 5723318
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, No. 5723318omi
; APPLICANT: Kojima, Tetsuo
; APPLICANT: Oh-Eda, Masayoshi
; APPLICANT: Hattori, Kunihiko
; TITLE OF INVENTION: Genes Coding for Megakaryocyte
; TITLE OF INVENTION: Potentiator

;; NUMBER OF SEQUENCES: 37
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Birch, Stewart, Kolasch & Birch
;; STREET: P.O. Box 747
;; CITY: Falls Church
;; STATE: Virginia
;; COUNTRY: USA
;; ZIP: 22040-0747
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/426,819A
;; FILING DATE: 21-APR-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Murphy Jr., Gerald M.
;; REGISTRATION NUMBER: 28,977
;; REFERENCE/DOCKET NUMBER: 230-107P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-205-8000
;; TELEFAX: 703-205-8050
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 16 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; HYPOTHEICAL: NO
;; FRAGMENT TYPE: N-terminal
;; IMMEDIATE SOURCE:
;; CLONE: HPCY5
;; FEATURE:
;; NAME/KEY: Peptide
;; LOCATION: 1..16
;; OTHER INFORMATION: /label= peptide
;; OTHER INFORMATION: /note= "amino terminal peptide of Meg POT, = sequence 1 in
;; US-08-426-819A-1

Query Match 100.0%; Score 19; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 9 AAPL 12

RESULT 35
US-08-426-819A-2
; Sequence 2, Application US/08426819A
; Patent No. 5723318
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, No. 5723318omi
; APPLICANT: Kojima, Tetsuo
; APPLICANT: Oh-Eda, Masayoshi
; APPLICANT: Hattori, Kunihiko
; TITLE OF INVENTION: Genes Coding for Megakaryocyte
; TITLE OF INVENTION: Potentiator
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,819A
; FILING DATE: 21-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 230-107P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; IMMEDIATE SOURCE:
; CLONE: HPCY5
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..16
; OTHER INFORMATION: /label= peptide
; OTHER INFORMATION: /note= "amino terminal sequence of MegPOT, = sequence 2 in Ta
; US-08-426-819A-2

Query Match 100.0%; Score 19; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 8 AAPL 11

RESULT 36
US-08-426-819A-3
; Sequence 3, Application US/08426819A
; Patent No. 5723318
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, No. 5723318omi
; APPLICANT: Kojima, Tetsuo
; APPLICANT: Oh-Eda, Masayoshi
; APPLICANT: Hattori, Kunihiko
; TITLE OF INVENTION: Genes Coding for Megakaryocyte
; TITLE OF INVENTION: Potentiator
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,819A
; FILING DATE: 21-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; 
```

```

; REFERENCE/DOCKET NUMBER: 230-107P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; IMMEDIATE SOURCE:
; CLONE: HPCY5
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..16
; OTHER INFORMATION: /label= peptide
; OTHER INFORMATION: /note= "amino-terminal sequence of MegPOT, = sequence 3 in
; US-08-426-819A-3

Query Match 100.0%; Score 19; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 7 AAPL 10

RESULT 37
US-08-735-963-4
; Sequence 4, Application US/08735963
; Patent No. 5804433
; GENERAL INFORMATION:
; APPLICANT: Gray A., Kevin
; APPLICANT: Childs D., John
; APPLICANT: Squires H., Charles
; TITLE OF INVENTION: A RHODOCOCCLUS FLAVIN REDUCTASE
; TITLE OF INVENTION: COMPLEMENTING DSZA AND DSZC ACTIVITY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/735,963
; FILING DATE: 23-OCT-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmore, Carolyn S.
; REGISTRATION NUMBER: 37,567
; REFERENCE/DOCKET NUMBER: EBC96-03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)861-6240
; TELEFAX: (617)861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; 
```

US-08-735-963-4

Query Match 100.0%; Score 19; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 11 AAPL 14

RESULT 38

US-09-105-057-4
; Sequence 4, Application US/09105057
; Patent No. 5919683
; GENERAL INFORMATION:
; APPLICANT: Gray, Kevin A.
; APPLICANT: Childs, John D.
; APPLICANT: Squires, Charles H.
; TITLE OF INVENTION: A RHODOCOCCLUS FLAVIN REDUCTASE
; TITLE OF INVENTION: COMPLEMENTING DSZA AND DSZC ACTIVITY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,057
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/735,963
; FILING DATE: 23-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmore, Carolyn S.
; REGISTRATION NUMBER: 37,567
; REFERENCE/DOCKET NUMBER: EBC96-032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)861-6240
; TELEFAX: (781)861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-057-4

Query Match 100.0%; Score 19; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 11 AAPL 14

RESULT 39
US-09-304-214-4
; Sequence 4, Application US/09304214
; Patent No. 6274372
; GENERAL INFORMATION:
; APPLICANT: Gray, Kevin A.
; APPLICANT: Childs, John D.

APPLICANT: Squires, Charles H.
; TITLE OF INVENTION: A RHODOCOCCLUS FLAVIN REDUCTASE
; TITLE OF INVENTION: COMPLEMENTING DSZA AND DSZC ACTIVITY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02421-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/304,214
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,057
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/735,963
; FILING DATE: 23-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmore, Carolyn S.
; REGISTRATION NUMBER: 37,567
; REFERENCE/DOCKET NUMBER: EBC96-0322
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)861-6240
; TELEFAX: (781)861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-304-214-4

Query Match 100.0%; Score 19; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 11 AAPL 14

RESULT 40
US-08-124-981A-26
; Sequence 26, Application US/08124981A
; Patent No. 5837840
; GENERAL INFORMATION:
; APPLICANT: Crabtree, Gerald R.
; APPLICANT: No. 5837840throp, Jeffrey P.
; TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/124,981A
; FILING DATE: 20-SEP-1993
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 5490A-226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-124-981A-26

```

```

Query Match      100.0%; Score 19; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 AAPL 4
        ||||
Db      11 AAPL 14

```

Search completed: December 6, 2002, 13:31:51
 Job time : 10.3333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:30:14 ; Search time 4.88889 Seconds
(without alignments)
13.289 Million cell updates/sec

Title: US-10-033-526-3

Perfect score: 19

Sequence: 1 AAPL 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*

2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*

10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*

12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*

13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|---------------------|-------------------|
| 1 | 19 | 100.0 | 4 | 10 | US-09-935-744-2 | Sequence 2, Appli |
| 2 | 19 | 100.0 | 4 | 12 | US-10-033-526-3 | Sequence 3, Appli |
| 3 | 19 | 100.0 | 9 | 10 | US-09-988-019-29 | Sequence 29, Appl |
| 4 | 19 | 100.0 | 9 | 10 | US-09-894-018-315 | Sequence 315, App |
| 5 | 19 | 100.0 | 15 | 10 | US-09-829-549A-31 | Sequence 31, Appl |
| 6 | 19 | 100.0 | 19 | 12 | US-10-001-879-133 | Sequence 133, App |
| 7 | 19 | 100.0 | 20 | 10 | US-09-813-333-46 | Sequence 46, Appl |
| 8 | 19 | 100.0 | 21 | 10 | US-09-813-333-37 | Sequence 37, Appl |
| 9 | 19 | 100.0 | 21 | 10 | US-09-813-333-53 | Sequence 53, Appl |
| 10 | 19 | 100.0 | 24 | 10 | US-09-864-761-37207 | Sequence 37207, A |
| 11 | 19 | 100.0 | 27 | 10 | US-09-060-878-3 | Sequence 3, Appli |
| 12 | 19 | 100.0 | 27 | 10 | US-09-864-761-40615 | Sequence 40615, A |
| 13 | 19 | 100.0 | 30 | 10 | US-09-864-761-34374 | Sequence 34374, A |
| 14 | 19 | 100.0 | 31 | 10 | US-09-864-761-43766 | Sequence 43766, A |
| 15 | 19 | 100.0 | 32 | 10 | US-09-821-883-11 | Sequence 11, Appl |
| 16 | 19 | 100.0 | 35 | 10 | US-09-864-761-33421 | Sequence 33421, A |
| 17 | 19 | 100.0 | 37 | 9 | US-10-108-915-10 | Sequence 10, Appl |
| 18 | 19 | 100.0 | 40 | 10 | US-09-864-761-46389 | Sequence 46389, A |
| 19 | 19 | 100.0 | 41 | 8 | US-08-424-5508-426 | Sequence 426, App |

| | | | | | | |
|----|----|-------|-----|----|---------------------|--------------------|
| 20 | 19 | 100.0 | 41 | 10 | US-09-864-761-41807 | Sequence 41807, A |
| 21 | 19 | 100.0 | 41 | 10 | US-09-864-761-42843 | Sequence 42843, A |
| 22 | 19 | 100.0 | 43 | 9 | US-09-975-143-5 | Sequence 5, Appli |
| 23 | 19 | 100.0 | 46 | 10 | US-09-864-761-36286 | Sequence 36286, A |
| 24 | 19 | 100.0 | 46 | 10 | US-09-864-761-48342 | Sequence 48342, A |
| 25 | 19 | 100.0 | 58 | 10 | US-09-864-761-35811 | Sequence 35811, A |
| 26 | 19 | 100.0 | 63 | 10 | US-09-764-869-833 | Sequence 833, App |
| 27 | 19 | 100.0 | 65 | 10 | US-09-864-761-33840 | Sequence 33840, A |
| 28 | 19 | 100.0 | 69 | 10 | US-09-291-809C-7 | Sequence 7, Appli |
| 29 | 19 | 100.0 | 69 | 10 | US-09-845-849-7 | Sequence 7, Appli |
| 30 | 19 | 100.0 | 69 | 10 | US-09-820-893-71 | Sequence 71, Appl |
| 31 | 19 | 100.0 | 72 | 10 | US-09-864-761-36594 | Sequence 36594, A |
| 32 | 19 | 100.0 | 73 | 9 | US-10-082-659-9 | Sequence 9, Appli |
| 33 | 19 | 100.0 | 80 | 10 | US-09-730-617-72 | Sequence 72, Appl |
| 34 | 19 | 100.0 | 81 | 10 | US-09-820-893-115 | Sequence 115, App |
| 35 | 19 | 100.0 | 85 | 10 | US-09-730-617-71 | Sequence 71, Appl |
| 36 | 19 | 100.0 | 85 | 10 | US-09-730-617-73 | Sequence 73, Appl |
| 37 | 19 | 100.0 | 86 | 10 | US-09-915-582-62 | Sequence 62, Appl |
| 38 | 19 | 100.0 | 90 | 10 | US-09-925-300-1019 | Sequence 1019, App |
| 39 | 19 | 100.0 | 93 | 10 | US-09-764-870-533 | Sequence 533, App |
| 40 | 19 | 100.0 | 94 | 10 | US-09-731-872-275 | Sequence 275, App |
| 41 | 19 | 100.0 | 96 | 9 | US-09-764-868-881 | Sequence 881, App |
| 42 | 19 | 100.0 | 101 | 10 | US-09-840-787-37 | Sequence 37, Appl |
| 43 | 19 | 100.0 | 104 | 10 | US-09-731-872-478 | Sequence 478, App |
| 44 | 19 | 100.0 | 111 | 10 | US-09-893-737-120 | Sequence 120, App |
| 45 | 19 | 100.0 | 111 | 10 | US-09-746-491-14 | Sequence 14, Appl |

ALIGNMENTS

RESULT 1

US-09-935-744-2

Sequence 2, Application US/09935744

Patent No. US20020137118A1

GENERAL INFORMATION:

APPLICANT: Inouye, Masayori

APPLICANT: Fu, Xuan

APPLICANT: Shinde, Ujwal

TITLE OF INVENTION: Biologically Active Protein Folding Intermediates

FILE REFERENCE: 266/223

CURRENT APPLICATION NUMBER: US/09/935,744

CURRENT FILING DATE: 2001-08-24

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 4

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: The sequence is part of a synthetic peptide that is used as a

OTHER INFORMATION: strate for determining the activation time of a stable crossli

NAME/KEY: MISC_FEATURE

LOCATION: (1)..(4)

OTHER INFORMATION: Sequence is preceded by N-succinyl and followed by p-nitroanil

US-09-935-744-2

Query Match 100.0%; Score 19; DB 10; Length 4;

Best Local Similarity 100.0%; Pred. No. 8.6e+04;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4

Db 1 AAPL 4

RESULT 2

US-10-033-526-3

Sequence 3, Application US/10033526

Patent No. US20020147999A1

GENERAL INFORMATION:

APPLICANT: Robert W. Mahley

APPLICANT: Yadong Huang
TITLE OF INVENTION: Methods of Treating Disorders Related to
TITLE OF INVENTION: APOE
FILE REFERENCE: UCAL217
CURRENT APPLICATION NUMBER: US/10/033,526
CURRENT FILING DATE: 2001-11-02
PRIOR APPLICATION NUMBER: 60/245,737
PRIOR FILING DATE: 2000-11-03
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 4
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
US-10-033-526-3

Query Match 100.0%; Score 19; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 1 AAPL 4

RESULT 3

US-09-988-019-29
Sequence 29, Application US/09988019
Patent No. US20020102277A1

GENERAL INFORMATION:

APPLICANT: SOKOL, Pamela A.

KOOL, Cora D.

TITLE OF INVENTION: CONSERVED METALLOPROTEASE EPITOPES

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: USA

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/988,019

FILING DATE: 16-NO. US20020102277A1-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/772,282

FILING DATE: 20-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: STANEX REA, Teresa

REGISTRATION NUMBER: 30,427

REFERENCE/DOCKET NUMBER: 024916-005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 29:

US-09-988-019-29

Query Match 100.0%; Score 19; DB 10; Length 9;

Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 4 AAPL 7

RESULT 4

US-09-894-018-315
Sequence 315, Application US/09894018
Patent No. US20020119127A1

GENERAL INFORMATION:

APPLICANT: EPIMUNE, Inc.

APPLICANT: Sette, Alessandro

APPLICANT: Chestnut, Robert

APPLICANT: Livingston, Brian

APPLICANT: Baker, Denisw

APPLICANT: Newman, Mark

APPLICANT: Brown, David

TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING

TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY

FILE REFERENCE: 39963-20033.00

CURRENT APPLICATION NUMBER: US/09/894,018

CURRENT FILING DATE: 2001-05-27

PRIOR APPLICATION NUMBER: PCT/US00/35568

PRIOR FILING DATE: 2000-12-28

PRIOR APPLICATION NUMBER: US 60/173,390

PRIOR FILING DATE: 1999-12-28

PRIOR APPLICATION NUMBER: US 60/284,221

PRIOR FILING DATE: 2001-04-16

NUMBER OF SEQ ID NOS: 368

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 315

LENGTH: 9

TYPE: PRT

ORGANISM: Transgenic mouse

US-09-894-018-315

Query Match 100.0%; Score 19; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
1111
Db 4 AAPL 7

RESULT 5

US-09-829-549A-31
Sequence 31, Application US/09829549A
Patent No. US20020052484A1

GENERAL INFORMATION:

APPLICANT: The Curators of the University of Missouri

TITLE OF INVENTION: PHAGE DISPLAY SELECTION OF ANTI FUNGAL PEPTIDES

FILE REFERENCE: UMO 1521.1

CURRENT APPLICATION NUMBER: US/09/829,549A

CURRENT FILING DATE: 2001-04-10

PRIOR APPLICATION NUMBER: US 60/195,785

PRIOR FILING DATE: 2000-04-10

NUMBER OF SEQ ID NOS: 48

SOFTWARE: PatentIn version 3.0

SEQ ID NO 31

LENGTH: 15

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: DOMAIN

LOCATION: (1)..(15)

OTHER INFORMATION: Random peptide insert

US-09-829-549A-31

Query Match 100.0%; Score 19; DB 10; Length 15;

Best Local Similarity 100.0%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 9 AAPL 12

RESULT 6

US-10-001-879-133
; Sequence 133, Application US/10001879
; Patent No. US20020127237A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Caferkey, Robert
; APPLICANT: Ali, Shujath
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0281
; CURRENT APPLICATION NUMBER: US/10/001,879
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,188
; PRIOR FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 133
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-879-133

Query Match 100.0%; Score 19; DB 12; Length 19;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 7 AAPL 10

RESULT 7

US-09-813-333-46
; Sequence 46, Application US/09813333
; Patent No. US20020119160A1
; GENERAL INFORMATION:
; APPLICANT: Degroot, Anne S
; TITLE OF INVENTION: Human T Cell Response to MHC-Binding Motif Clusters
; FILE REFERENCE: 17999-004 US
; CURRENT APPLICATION NUMBER: US/09/813,333
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,834
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 46
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-813-333-46

Query Match 100.0%; Score 19; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 12 AAPL 15

RESULT 8

US-09-813-333-37
; Sequence 37, Application US/09813333
; Patent No. US20020119160A1
; GENERAL INFORMATION:
; APPLICANT: Degroot, Anne S
; TITLE OF INVENTION: Human T Cell Response to MHC-Binding Motif Clusters
; FILE REFERENCE: 17999-004 US
; CURRENT APPLICATION NUMBER: US/09/813,333
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,834
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 37
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-813-333-37

Query Match 100.0%; Score 19; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 13 AAPL 16

RESULT 9

US-09-813-333-53
; Sequence 53, Application US/09813333
; Patent No. US20020119160A1
; GENERAL INFORMATION:
; APPLICANT: Degroot, Anne S
; TITLE OF INVENTION: Human T Cell Response to MHC-Binding Motif Clusters
; FILE REFERENCE: 17999-004 US
; CURRENT APPLICATION NUMBER: US/09/813,333
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,834
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 53
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-813-333-53

Query Match 100.0%; Score 19; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 12 AAPL 15

RESULT 10

US-09-864-761-37207
; Sequence 37207, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04

```
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37207
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AEO00659.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 9.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 18
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.4
US-09-864-761-37207

Query Match      100.0%; Score 19; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 AAPL 4
Db 17 AAPL 20

```
RESULT 11
US-09-060-878-3
; Sequence 3, Application US/09060878A
; Patent No. US20010006642A1
; GENERAL INFORMATION:
; APPLICANT: Steidler, Lothar
; APPLICANT: Remaut, Erik
; APPLICANT: Wells, Jeremy Mark
; APPLICANT: Le Page, Richard William
; TITLE OF INVENTION: Delivery of Biologically Active
```

```
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 8865-009-999
; CURRENT APPLICATION NUMBER: US/09/060,878A
; CURRENT FILING DATE: 1998-04-16
; EARLIER APPLICATION NUMBER: GB9521568.7
; EARLIER FILING DATE: 1995-10-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Leader Sequence of Usp45 in Lactococcus lactis
US-09-060-878-3

Query Match      100.0%; Score 19; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 AAPL 4
Db 19 AAPL 22

```
RESULT 12
US-09-864-761-40615
; Sequence 40615, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
```

```
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40615
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC021498.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4
; US-09-864-761-40615

Query Match
Best Local Similarity 100.0%; Score 19; DB 10; Length 27;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 1 AAPL 4

RESULT 13
; US-09-864-761-34374
; Sequence 34374, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34374
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007914.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9.3
; NAME/KEY: unsure
; LOCATION: 22
; US-09-864-761-34374

Query Match
Best Local Similarity 100.0%; Score 19; DB 10; Length 30;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 25 AAPL 28

RESULT 14
; US-09-864-761-43766
; Sequence 43766, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43766
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005959.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.71
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.64
; OTHER INFORMATION: EST_HUMAN HIT: AI221385.1, EVALUATE 2.00e-11
US-09-864-761-43766
```

```
Query Match          100.0%; Score 19; DB 10; Length 31;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPL 4
        ||||
Db       24 AAPL 27
```

```
RESULT 15
US-09-821-883-11
; Sequence 11, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-821-883-11
```

```
Query Match          100.0%; Score 19; DB 10; Length 32;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPL 4
        ||||
Db       3 AAPL 6
```

RESULT 16

US-09-864-761-33421

```
; Sequence 33421, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33421
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007002.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EST_HUMAN HIT: AA744947.1, EVALUATE 1.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P27197, EVALUATE 2.80e+00
US-09-864-761-33421
```

```
Query Match          100.0%; Score 19; DB 10; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
```


Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APL 4
|||||
Db 27 APL 30

RESULT 17
US-10-108-915-10
; Sequence 10, Application US/10108915
; Patent No. US20020177204A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Shen, Jennie
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: Geranylgeranyl Pyrophosphate Synthases
; FILE REFERENCE: B01286 US NA
; CURRENT APPLICATION NUMBER: US/10/108,915
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/452,238
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,592
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-december-02
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Zea mays
US-10-108-915-10

Query Match 100.0%; Score 19; DB 9; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APL 4
|||||
Db 22 APL 25

RESULT 18
US-09-864-761-46389
; Sequence 46389, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46389
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION: MAP TO AC011505.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EST_HUMAN HIT: B0185826.1, EVALUATE 1.90e+00
; OTHER INFORMATION: SWISSPROT HIT: 008808, EVALUATE 7.30e-01
US-09-864-761-46389

Query Match 100.0%; Score 19; DB 10; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APL 4
|||||
Db 11 APL 14

RESULT 19
US-08-424-550B-426
; Sequence 426, Application US/08424550B
; Patent No. US20020119447A1
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMM J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHRHOF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUTIK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```
; APPLICATION NUMBER: US/08/424,550B
; FILING DATE:
; CLASSIFICATION: 435435
; ATTORNEY/AGENT INFORMATION:
; NAME: POREBSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 426:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 41 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-424-550B-426

Query Match          100.0%; Score 19; DB 8; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPL 4
        ||||
Db       8 AAPL 11

RESULT 20
US-09-864-761-41807
; Sequence 41807, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
```

```
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41807
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006549.28
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.1
; OTHER INFORMATION: SWISSPROT HIT: Q9X3X1, EVALUATE 2.80e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF568063.1, EVALUATE 6.00e-16
; US-09-864-761-41807

Query Match          100.0%; Score 19; DB 10; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPL 4
        ||||
Db       21 AAPL 24

RESULT 21
US-09-864-761-42843
; Sequence 42843, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42843
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121657.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EST HUMAN HIT: A1673575.1, EVALUE 2.00e-11
; OTHER INFORMATION: SWISSPROT HIT: Q62168, EVALUE 1.60e+00
US-09-864-761-42843

Query Match          100.0%; Score 19; DB 10; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPL 4
       1111
Db      16 AAPL 19

RESULT 22
US-09-975-143-5
; Sequence 5, Application US/09975143
; Patent No. US20020155513A1
; GENERAL INFORMATION:
; APPLICANT: HSU, Daniel, K.
; APPLICANT: LIU, Fu-Tong
; APPLICANT: DOWLING, Christopher, A.
; TITLE OF INVENTION: GALECTIN EXPRESSION IS INDUCED IN
; TITLE OF INVENTION: CIRRHOTIC LIVER AND HEPATOCELLULAR CARCINOMA
; FILE REFERENCE: DANHSU.001C1
; CURRENT APPLICATION NUMBER: US/09/975,143
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08561
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 43
; TYPE: PRT
; ORGANISM: chicken
US-09-975-143-5

Query Match          100.0%; Score 19; DB 9; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPL 4
       1111
Db      22 AAPL 25

RESULT 23
US-09-864-761-36286
; Sequence 36286, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
```

```

; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36286
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL022313.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 54
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 9.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.6
US-09-864-761-36286

Query Match          100.0%; Score 19; DB 10; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPL 4
       1111
Db      24 AAPL 27

RESULT 24
US-09-864-761-48342
```

```
; Sequence 48342, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48342
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011338.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.8
; OTHER INFORMATION: SWISSPROT HIT: Q03603, EVALU6 6.20e+00
US-09-864-761-48342

Query Match      100.0%; Score 19; DB 10; Length 46;
Best Local Similarity 100.0%; Pred. NO. 1.5e+02;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35811
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007378.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 9.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
; OTHER INFORMATION: EST_HUMAN HIT: BF334188.1, EVALU6 8.20e-02
US-09-864-761-35811

Query Match      100.0%; Score 19; DB 10; Length 58;
Best Local Similarity 100.0%; Pred. NO. 1.9e+02;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 36 AAPL 39

RESULT 26

US-09-764-869-833

; Sequence 833, Application US/09764869

; Patent No. US20020061521A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC007

; CURRENT APPLICATION NUMBER: US/09/764,869

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 2442

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 833

; LENGTH: 63

; TYPE: PRT

; ORGANISM: Homo sapiens

; NAME/KEY: SITE

; LOCATION: (50)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (55)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (58)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-764-869-833

Query Match

Best Local Similarity 100.0%; Score 19; DB 10; Length 63;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4

Db 16 AAPL 19

RESULT 27

US-09-864-761-33840

; Sequence 33840, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aecomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 33840

; LENGTH: 65

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC007790.1

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3

; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9

US-09-864-761-33840

Query Match

Best Local Similarity 100.0%; Score 19; DB 10; Length 65;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4

Db 42 AAPL 45

RESULT 28

US-09-291-809C-7

; Sequence 7, Application US/09291809C

; Patent No. US20010049831A1

; GENERAL INFORMATION:

; APPLICANT: detlef weigel

; APPLICANT: Salk Institute

; TITLE OF INVENTION: FLOWERING LOCUS T (FT) AND GENETICALLY

; FILE REFERENCE: SALKINS.026CP1

; CURRENT APPLICATION NUMBER: US/09/291,809C

; CURRENT FILING DATE: 1999-04-14

; PRIOR APPLICATION NUMBER: PCT/US99/08151

; PRIOR FILING DATE: 1999-04-13

; PRIOR APPLICATION NUMBER: 09/060,726

; PRIOR FILING DATE: 1998-04-15

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 69

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-291-809C-7

Query Match

Best Local Similarity 100.0%; Score 19; DB 10; Length 69;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 2 AAPL 5

RESULT 29

US-09-845-849-7
; Sequence 7, Application US/09845849
; Patent No. US20020029395A1
; GENERAL INFORMATION:
; APPLICANT: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES
; APPLICANT: WEIGEL, Detlef
; APPLICANT: KARDAILSKY, Igor
; TITLE OF INVENTION: FLOWERING LOCUS T (FT) AND GENETICALLY
; TITLE OF INVENTION: MODIFIED PLANTS HAVING MODULATED FLOWER DEVELOPMENT
; FILE REFERENCE: SALKINS.026DV1
; CURRENT APPLICATION NUMBER: US/09/845,849
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/060,726
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-845-849-7

Query Match 100.0%; Score 19; DB 10; Length 69;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 2 AAPL 5

RESULT 30

US-09-820-893-71
; Sequence 71, Application US/09820893
; Patent No. US20020076705A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 31 Human Secreted Proteins
; FILE REFERENCE: P2033P1
; CURRENT APPLICATION NUMBER: US/09/820,893
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/531,119
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/102,895
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (69)
; OTHER INFORMATION: Xaa equals stop translation
US-09-820-893-71

Query Match 100.0%; Score 19; DB 10; Length 69;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 2 AAPL 5

RESULT 31

US-09-864-761-36594
; Sequence 36594, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36594
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007336.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.4
; OTHER INFORMATION: EST_HUMAN HIT: BE261817.1, EVALUUE 7.00e+00
US-09-864-761-36594

Query Match 100.0%; Score 19; DB 10; Length 72;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0;

OY 1 AAPL 4
Db 28 AAPL 31

RESULT 32

US-10-082-659-9
; Sequence 9, Application US/10082659
; Patent No. US20020168783A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Cohen, Maurice
; APPLICANT: Colpitts, Tracey L.
; APPLICANT: Friedman, Paula N.
; APPLICANT: Granados, Edward N.
; APPLICANT: Klass, Michael R.
; APPLICANT: Russell, John C.
; APPLICANT: Stroupe, Stephen D.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR
; FILE REFERENCE: 6171 US.P1
; CURRENT APPLICATION NUMBER: US/10/082,659
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/566,876
; PRIOR FILING DATE: 2000-05-08
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-082-659-9

Query Match 100.0%; Score 19; DB 9; Length 73;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 23 AAPL 26

RESULT 33

US-09-730-617-72
; Sequence 72, Application US/09730617
; Patent No. US20020068279A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine E
; APPLICANT: Prayaga, Sudhirdas K
; APPLICANT: Shinkets, Richard A
; APPLICANT: Rastelli, Luca
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20020068279A1 Proteins and Nucleic Acids Encoding the Sam
; FILE REFERENCE: 15966-609
; CURRENT APPLICATION NUMBER: US/09/730,617
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: 60/169,056
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 60/169,886
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/169,866
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/170,252
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/175,740
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 72

; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-730-617-72

Query Match 100.0%; Score 19; DB 10; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 23 AAPL 26

RESULT 34

US-09-820-893-115
; Sequence 115, Application US/09820893
; Patent No. US20020076705A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 31 Human Secreted Proteins
; FILE REFERENCE: P2033P1
; CURRENT APPLICATION NUMBER: US/09/820,893
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/531,119
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/102,895
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 115
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-820-893-115

Query Match 100.0%; Score 19; DB 10; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 13 AAPL 16

RESULT 35

US-09-730-617-71
; Sequence 71, Application US/09730617
; Patent No. US20020068279A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine E
; APPLICANT: Prayaga, Sudhirdas K
; APPLICANT: Shinkets, Richard A
; APPLICANT: Rastelli, Luca
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20020068279A1 Proteins and Nucleic Acids Encoding the
; FILE REFERENCE: 15966-609
; CURRENT APPLICATION NUMBER: US/09/730,617
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: 60/169,056
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 60/169,886
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/169,866
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/170,252
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/175,740
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 71

```
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-730-617-71

Query Match          100.0%; Score 19; DB 10; Length 85;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPL 4
        ||||
Db       23 AAPL 26

RESULT 36
US-09-730-617-73
; Sequence 73, Application US/09730617
; Patent No. US20020068279A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine E
; APPLICANT: Prayaga, Sudhirdas K
; APPLICANT: Shimkets, Richard A
; APPLICANT: Rastelli, Luca
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20020068279A1e1 Proteins and Nucleic Acids Encoding the San
; FILE REFERENCE: 15966-609
; CURRENT APPLICATION NUMBER: US/09/730,617
; CURRENT FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: 60/169,056
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 60/169,886
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/169,866
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/170,252
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/175,740
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-730-617-73

Query Match          100.0%; Score 19; DB 10; Length 85;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPL 4
        ||||
Db       23 AAPL 26

RESULT 37
US-09-915-582-62
; Sequence 62, Application US/09915582
; Patent No. US20020120103A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: PS723P1
; CURRENT APPLICATION NUMBER: US/09/915,582
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
```

```
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-915-582-62

Query Match          100.0%; Score 19; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPL 4
        ||||
Db       33 AAPL 36

RESULT 38
US-09-925-300-1019
; Sequence 1019, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1019
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1019

Query Match          100.0%; Score 19; DB 10; Length 90;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAPL 4
        ||||
Db       66 AAPL 69

RESULT 39
US-09-764-870-533
; Sequence 533, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 533
; LENGTH: 93
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (53)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (92)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-870-533
```

```
Query Match          100.0%; Score 19; DB 10; Length 93;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPL 4
        |||
Db       71 AAPL 74
```

```
RESULT 40
US-09-731-872-275
; Sequence 275, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouquelere, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 275
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -25..-1
US-09-731-872-275
```

```
Query Match          100.0%; Score 19; DB 10; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAPL 4
        |||
Db       86 AAPL 89
```

```
Search completed: December 6, 2002, 13:42:06
Job time : 4.88889 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:25:59 ; Search time 10.2222 Seconds
(without alignments)
37.618 Million cell updates/sec

Title: US-10-033-526-3
Perfect score: 19
Sequence: 1 AAPL 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 19 | 100.0 | 18 | 2 A49857 | megakaryocyte pote |
| 2 | 19 | 100.0 | 30 | 2 G71359 | hypothetical prote |
| 3 | 19 | 100.0 | 32 | 2 T36974 | hypothetical prote |
| 4 | 19 | 100.0 | 34 | 2 PC1267 | subtilisin inhibit |
| 5 | 19 | 100.0 | 36 | 2 S78239 | ycf32 protein - Od |
| 6 | 19 | 100.0 | 38 | 2 T01677 | pseudo-kallikrein |
| 7 | 19 | 100.0 | 45 | 2 PN0592 | tyrosine 3-monooxy |
| 8 | 19 | 100.0 | 53 | 2 E71270 | hypothetical prote |
| 9 | 19 | 100.0 | 60 | 2 C82812 | hypothetical prote |
| 10 | 19 | 100.0 | 62 | 2 G70988 | hypothetical prote |
| 11 | 19 | 100.0 | 62 | 2 E82630 | hypothetical prote |
| 12 | 19 | 100.0 | 63 | 2 B81304 | hypothetical prote |
| 13 | 19 | 100.0 | 66 | 2 H87276 | probable periplasm |
| 14 | 19 | 100.0 | 67 | 2 E59147 | hypothetical prote |
| 15 | 19 | 100.0 | 67 | 2 AB0396 | conotoxin Tx5.2 pr |
| 16 | 19 | 100.0 | 72 | 2 B81922 | hypothetical prote |
| 17 | 19 | 100.0 | 75 | 2 B82758 | probable integral |
| 18 | 19 | 100.0 | 76 | 2 A59043 | hypothetical prote |
| 19 | 19 | 100.0 | 76 | 2 G98238 | contulakin-G precu |
| 20 | 19 | 100.0 | 77 | 2 T09231 | probable peroxidas |
| 21 | 19 | 100.0 | 78 | 2 AE0572 | hypothetical secre |
| 22 | 19 | 100.0 | 80 | 2 T24150 | hypothetical prote |
| 23 | 19 | 100.0 | 80 | 2 B95851 | hypothetical prote |
| 24 | 19 | 100.0 | 80 | 2 AF2836 | hypothetical prote |
| 25 | 19 | 100.0 | 82 | 2 G01978 | calmodulin-depende |
| 26 | 19 | 100.0 | 83 | 2 C81840 | hypothetical prote |
| 27 | 19 | 100.0 | 83 | 2 T35979 | hypothetical prote |
| 28 | 19 | 100.0 | 84 | 2 A13550 | hypothetical prote |
| 29 | 19 | 100.0 | 85 | 2 A39747 | hypothetical prote |

| | | | | | |
|----|----|-------|----|----------|--------------------|
| 30 | 19 | 100.0 | 86 | 2 C90666 | probable DNA bindi |
| 31 | 19 | 100.0 | 86 | 2 G85516 | probable DNA bindi |
| 32 | 19 | 100.0 | 88 | 2 PC4061 | pyrrolo-quinoline- |
| 33 | 19 | 100.0 | 89 | 1 C36869 | probable export pr |
| 34 | 19 | 100.0 | 89 | 2 S78699 | flagellar biosynth |
| 35 | 19 | 100.0 | 89 | 2 AE0753 | flagellar biosynth |
| 36 | 19 | 100.0 | 89 | 2 A10221 | flagellar biosynth |
| 37 | 19 | 100.0 | 89 | 2 H90964 | probable export pr |
| 38 | 19 | 100.0 | 89 | 2 H85812 | flagellar biosynth |
| 39 | 19 | 100.0 | 90 | 2 T31980 | hypothetical prote |
| 40 | 19 | 100.0 | 91 | 1 H69198 | hypothetical prote |
| 41 | 19 | 100.0 | 91 | 2 G01975 | calmodulin-depende |
| 42 | 19 | 100.0 | 92 | 2 S41662 | L-phenylalanine ox |
| 43 | 19 | 100.0 | 93 | 2 E87678 | hypothetical prote |
| 44 | 19 | 100.0 | 93 | 2 H75502 | conserved hypothet |
| 45 | 19 | 100.0 | 94 | 2 G81047 | hypothetical prote |

ALIGNMENTS

RESULT 1
A49857
megakaryocyte potentiating factor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995
C:Accession: A49857
R:Yamaguchi, N.; Hattori, K.; Oh-eda, M.; Kojima, T.; Imai, N.; Ochi, N.
J. Biol. Chem. 269, 805-808, 1994
A:Title: A novel cytokine exhibiting megakaryocyte potentiating activity from a human
A:Reference number: A49857; MUID:94117486; PMID:8288629
A:Accession: A49857
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-18 <YAM>
A:Experimental source: pancreatic tumor cell line HPC-Y5
A>Note: sequence extracted from NCBI backbone (NCBIP:142670)

Query Match 100.0%; Score 19; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
DB 9 AAPL 12

RESULT 2
G71359
hypothetical protein TP0161 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C:Accession: G71359
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G
rison, J.; Khailak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: G71359
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-30 <COL>
A:Cross-references: GB:AE001200; GB:AE000520; NID:g3322419; PIDN:AAC65153.1; PID:g3332
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0161

Query Match 100.0%; Score 19; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4

Db 6 AAPL 9

RESULT 3

T36974

hypothetical protein SCJ11.03 - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T36974

R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A:Reference number: 221618

A:Accession: T36974

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-32 <OLI>

A:Cross-references: EMBL:AL109949; PIDN:CAB52888.1; GSPDB:GN00070; SCOEDB:SCJ11.03

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SCJ11.03

Query Match

100.0%; Score 19; DB 2; Length 32;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4

||||

Db 5 AAPL 8

RESULT 4

PC1267

subtilisin inhibitor 3 - Streptomyces coelicolor (fragment)

C:Species: Streptomyces coelicolor

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PC1267

R:Taguchi, S.; Kikuchi, H.; Kojima, S.; Kumagai, I.; Nakase, T.; Miura, K.; Momose, H.

Biosci. Biotechnol. Biochem. 57, 522-524, 1993

A:Title: High frequency of SSI-like protease inhibitors among Streptomyces.

A:Reference number: PC1260; MUID:93222542; PMID:7763545

A:Accession: PC1267

A:Molecule type: protein

A:Residues: 1-34 <TAG>

A:Experimental source: strain KCC-S006

C:Superfamily: plasminostreptin

C:Keywords: serine proteinase inhibitor

Query Match

100.0%; Score 19; DB 2; Length 34;

Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4

||||

Db 19 AAPL 22

RESULT 5

S78239

ycf32 protein - Odontella sinensis chloroplast

C:Species: Chloroplast Odontella sinensis

C:Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 20-Jun-2000

C:Accession: S78239; S78293

R:Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.

Plant Mol. Biol. Rep. 13, 336-342, 1995

A:Title: The Chloroplast Genome of a chlorophyll a+c-containing Alga, Odontella sinensis

A:Reference number: S78238

A:Accession: S78239

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-36 <KOW>

A:Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91612.1; PID:g1185129

A:Genetics: 5GEN

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

A:Note: 5' gene

A:Accession: S78293

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-36 <KOF>

A:Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91666.1; PID:g1185183

A:Genetics: 3GEN

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

A:Note: 3' gene (C)

C:Genetics: <GEN1>

A:Gene: ycf32

A:Genome: chloroplast

A:Note: gene located on inverted repeat IRA

C:Genetics: <GEN2>

A:Gene: ycf32'

A:Genome: chloroplast

A:Note: gene located on inverted repeat IRb

C:Superfamily: hypothetical protein ycf32

C:Keywords: chloroplast

Query Match

100.0%; Score 19; DB 2; Length 36;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4

||||

Db 9 AAPL 12

RESULT 6

T01677

pseudo-kallikrein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 20-Jun-2000

C:Accession: T01677

R:Evans, B.A.; Drinkwater, C.C.; Richards, R.I.

J. Biol. Chem. 262, 8027-8034, 1987

A:Title: Mouse glandular kallikrein genes: Structure and partial sequence analysis of

A:Reference number: I55260; MUID:87250386; PMID:3036794

A:Accession: T01677

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-38 <EVA>

A:Cross-references: EMBL:M18581; NID:g198496

C:Superfamily: trypsin; trypsin homology

Query Match

100.0%; Score 19; DB 2; Length 38;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4

||||

Db 1 AAPL 4

RESULT 7

PN0592

tyrosine 3-monooxygenase (EC 1.14.16.2) type 1 and 2 - common gibbon (fragment)

N:Alternate names: oxygen oxidoreductase; tetrahydropteridine; tyrosine hydroxylase

C:Species: Hylobates lar (common gibbon, white-handed gibbon)

C:Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 31-Mar-2000

C:Accession: PN0592

R:Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.

Biochem. Biophys. Res. Commun. 195, 158-165, 1993

A:Title: Increased heterogeneity of tyrosine hydroxylase in humans.

A:Reference number: PN0575; MUID:93371398; PMID:7689834

A:Accession: PN0592

A:Molecule type: genomic RNA

A:Residues: 1-45 <ICH>

A:Cross-references: GB:L14794

A:Experimental source: lymph nodes

C:Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine bi

C:Superfamily: phenylalanine 4-monooxygenase
C:Keywords: biotpterin; monooxygenase; oxidoreductase

| | | | | |
|-----------------------|-----------------|--------------------|-----------|------------|
| Query Match | 100.0%; | Score 19; | DB 2; | Length 45; |
| Best Local Similarity | 100.0%; | Pred. No. 3.7e+02; | | |
| Matches 4; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

QY 1 AAPL 4
||||
Db 29 AAPL 32

RESULT 8

E71270
hypothetical protein TP0871 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C:Accession: E71270
R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: E71270
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-53 <COL>
A:Cross-references: GB:AE001257; GB:AE000520; NID:g3323182; PIDN:AAC65846.1; PID:g332319
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0871

| | | | | |
|-----------------------|-----------------|--------------------|-----------|------------|
| Query Match | 100.0%; | Score 19; | DB 2; | Length 53; |
| Best Local Similarity | 100.0%; | Pred. No. 3.2e+02; | | |
| Matches 4; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

QY 1 AAPL 4
||||
Db 42 AAPL 45

RESULT 9

C82812
hypothetical protein XF0382 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: C82812
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: C82812
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-60 <SIM>
A:Cross-references: GB:AE003890; GB:AE003849; NID:g9105215; PIDN:AAF83192.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincant, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation

C:Genetics:
A:Gene: XF0382

| | | | | |
|-----------------------|-----------------|--------------------|-----------|------------|
| Query Match | 100.0%; | Score 19; | DB 2; | Length 60; |
| Best Local Similarity | 100.0%; | Pred. No. 3.6e+02; | | |
| Matches 4; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

QY 1 AAPL 4
||||
Db 34 AAPL 37

RESULT 10

G70988
hypothetical protein RV1766 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: G70988
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: G70988
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-62 <COL>
A:Cross-references: GB:Z95890; GB:AL123456; NID:g3242245; PIDN:CAB09309.1; PID:e31901
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV1766

| | | | | |
|-----------------------|-----------------|--------------------|-----------|------------|
| Query Match | 100.0%; | Score 19; | DB 2; | Length 62; |
| Best Local Similarity | 100.0%; | Pred. No. 3.7e+02; | | |
| Matches 4; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

QY 1 AAPL 4
||||
Db 46 AAPL 49

RESULT 11

E82630
hypothetical protein XF1853 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: E82630
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: E82630
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-62 <SIM>
A:Cross-references: GB:AE004006; GB:AE003849; NID:g9106932; PIDN:AAF84659.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincant, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328

A:Contents: annotation
C:Genetics:
A:Gene: XP1853

Query Match 100.0%; Score 19; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. NO. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
||||
Db 6 AAPL 9

RESULT 12
B81304
Probable periplasmic protein Cj1021c [imported] - Campylobacter jejuni (strain NCTC 1116
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: B81304
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: B81304
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-63 <PAR>
A:Cross-references: GB:AL139077; GB:AL111168; NID:g6968444; PIDN:CAB73277.1; PID:g696845
C:Genetics:
A:Experimental source: serotype O2, strain NCTC 11168
A:Gene: Cj1021c

Query Match 100.0%; Score 19; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. NO. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
||||
Db 17 AAPL 20

RESULT 13
H87276
hypothetical protein CC0225 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87276
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87276
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-66 <STO>
A:Cross-references: GB:AE005673; NID:g13421354; PIDN:AAK22212.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0225

Query Match 100.0%; Score 19; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. NO. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
||||
Db 11 AAPL 14

RESULT 14
E59147

conotoxin Tx5.2 precursor [validated] - cone shell (Conus textile)
N:Alternate names: epsilon-conotoxin Txix
N:Contains: conotoxin Tx5a
C:Species: Conus textile (cloth-of-gold cone)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Oct-2000
C:Accession: E59147; A59044

R:Walker, C.S.; Steel, D.; Jacobsen, R.B.; Lirazan, M.B.; Cruz, L.J.; Hooper, D.; She
J. Biol. Chem. 274, 30664-30671, 1999
A:Title: The T-superfamily of conotoxins.
A:Reference number: A59147; MUID:99452958; PMID:10521453
A:Accession: E59147

A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA; protein
A:Residues: 1-67 <WAL>

A:Cross-references: GB:AF167167; NID:g6103612; PIDN:AAF03687.1; PID:g6103613
A:Note: submitted to GenBank, July 1999; named tx5a in GenBank entry AF167167, releas
R:Rigby, A.C.; Lucas-Meunier, E.; Kalume, D.E.; Czerwicz, E.; Hambe, B.; Dahlqvist, I
Proc. Natl. Acad. Sci. U.S.A. 96, 5758-5763, 1999
A:Title: A conotoxin from Conus textile with unusual posttranslational modifications
A:Reference number: A59044; MUID:99254114; PMID:10318957
A:Accession: A59044

A:Status: preliminary
A:Molecule type: protein
A:Residues: 51-63 <RIG>

C:Comment: This conotoxin reduces both presynaptic calcium influx and acetylcholine r
C:Superfamily: unassigned conotoxins
C:Keywords: bromine; carboxyglutamic acid; glycoprotein; hydroxyproline; toxin; venom
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-50/Domain: amino-terminal propeptide #status predicted <PRO>
F:51-63/Product: conotoxin tx5a #status experimental <MAT>
F:64-67/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:51,54/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental
F:52-58,53-59/Disulfide bonds: #status experimental
F:57/Modified site: 6-bromotyptophan (Trp) #status experimental
F:60/Binding site: carbohydrate (Thr) (covalent) #status experimental
F:63/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match 100.0%; Score 19; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. NO. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
||||
Db 61 AAPL 64

RESULT 15
AB0396
hypothetical protein YPO3260 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AB0396
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0396
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-67 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92494.1; PID:g15981194; GSPDB:GN00175
C:Genetics:
A:Gene: YPO3260

Query Match 100.0%; Score 19; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. NO. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
||||
Db 47 AAPL 50

RESULT 16

probable integral membrane protein NMA0779 [imported] - Neisseria meningitidis (strain Z
D81922
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: D81922
R.Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: D81922
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-72 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84062.1; PID:g737950
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0779

Query Match

Best Local Similarity 100.0%; Score 19; DB 2; Length 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||

Db 19 AAPL 22

RESULT 17

B82758
hypothetical protein XF0829 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: B82758
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82758
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-75 <SIM>
A:Cross-references: GB:AE003922; GB:AE003849; NID:g9105722; PIDN:AAE83639.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Relnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Brlones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0829

Query Match

Best Local Similarity 100.0%; Score 19; DB 2; Length 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||

Db 58 AAPL 61

RESULT 18

A59043
contulakin-G precursor - cone shell (Conus geographus)
C:Species: Conus geographus (geography cone)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 21-Jul-2000
C:Accession: A59043; B59043
R:Craig, A.G.; Norberg, T.; Griffin, D.; Hoeger, C.; Akhtar, M.; Schmidt, K.; Low, W.
a, B.M.
J. Biol. Chem. 274, 13752-13759, 1999
A:Title: Contulakin-G, an O-glycosylated invertebrate neurotensin.
A:Reference number: A59043; MUID:99253920; PMID:10318778
A:Accession: A59043
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-76 <CRA1>
A:Cross-references: GB:AF121108; NID:g4808952; PIDN:AAD30031.1; PID:g4808953
A:Accession: B59043
A:Molecule type: protein
A:Residues: 51-66 <CRA2>
C:Keywords: blocked amino end; glycoprotein; neurotoxin; pyroglutamic acid; venom
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-50/Domain: amino-terminal propeptide #status predicted <PRO>
F:51-66/Product: contulakin-G #status experimental <MAT>
F:67-76/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:51/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F:60/Binding site: carbohydrate (Thr) (covalent) #status experimental

Query Match

Best Local Similarity 100.0%; Score 19; DB 2; Length 76;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||

Db 16 AAPL 19

RESULT 19

G98238
hypothetical protein AGR_L1719 [imported] - Agrobacterium tumefaciens (strain C58, C
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: G98238
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: G98238
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-76 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK89433.1; PID:g15159294; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L1719
A:Map position: linear chromosome

Query Match

Best Local Similarity 100.0%; Score 19; DB 2; Length 76;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||

Db 9 AAPL 12

RESULT 20

T09231
probable peroxidase (EC 1.11.1.7) precursor prx10 - spinach (fragment)
C:Species: Spinacia oleracea (spinach)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 20-Jun-2000
C:Accession: T09231
R:Diogon, T.; Greppin, H.; Simon, P.

submitted to the EMBL Data Library, February 1998
A:Description: cDNA cloning and expression of various peroxidase genes in *Spinacia oleracea*
A:Reference number: Z16615
A:Accession: T09231
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-77 <DIO>
A:Cross-references: EMBL:Y16777
A:Experimental source: cultivar Nobel; leaf
C:Genetics:
A:Gene: prx10
C:Superfamily: peroxidase
C:Keywords: oxidoreductase
F:1-26/Domain: signal sequence
F:27-77/Product: peroxidase prx10 (fragment) #status predicted <SIG>
F:27-77/Product: peroxidase prx10 (fragment) #status predicted <MAT>

Query Match 100.0%; Score 19; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
| | | |
Db 59 AAPL 62

RESULT 21
AE0572
hypothetical secreted protein STY0613 [imported] - *Salmonella enterica* subsp. *enterica* s
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A:Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AE0572
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AE0572
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-78 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05047.1; PID:g16501823; GSPDB:GN00176
C:Genetics:
A:Gene: STY0613

Query Match 100.0%; Score 19; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
| | | |
Db 13 AAPL 16

RESULT 22
T24150
hypothetical protein R10H10.4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24150
R:Bardill, S.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19846
A:Accession: T24150
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-80 <WIL>
A:Cross-references: EMBL:Z70686; PIDN:CAA94608.1; GSPDB:GN00022; CESP:R10H10.4
A:Experimental source: clone R10H10
C:Genetics:
A:Gene: CESP:R10H10.4

A:Map position: 4
A:Introns: 48/3

Query Match 100.0%; Score 19; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
| | | |
Db 12 AAPL 15

RESULT 23
B95851
hypothetical protein [imported] - *Sinorhizobium meliloti* (strain 1021) megaplasmid ps
C:Species: *Sinorhizobium meliloti*
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: B95851
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her
proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: B95851
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-80 <KUR>
A:Cross-references: GB:AL591985; PIDN:CAC48474.1; PID:g15139946; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen,
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMB20074
A:Genome: plasmid

Query Match 100.0%; Score 19; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
| | | |
Db 9 AAPL 12

RESULT 24
AF2836
hypothetical protein Atu2118 [imported] - *Agrobacterium tumefaciens* (strain C58, Dupo
C:Species: *Agrobacterium tumefaciens*
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AF2836
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AF2836
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-80 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL43108.1; PID:g17740580; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu2118
A:Map position: circular chromosome

Query Match 100.0%; Score 19; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
DB 48 AAPL 51

RESULT 25

G01978

calmodulin-dependent protein kinase II gamma F isoform - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 18-Jun-1999

C:Accession: G01978

R:Kwiatkowski, A.P.; McGill, J.M.

submitted to the EMBL Data Library, July 1995

A:Reference number: G08974

A:Accession: G01978

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-82 <KWI>

A:Cross-references: EMBL:U32509; NID:g975884; PIDN:AAA75203.1; PID:g975885

C:Superfamily: Ca2+/calmodulin-dependent protein kinase II; protein kinase homology

Query Match

100.0%; Score 19; DB 2; Length 82;

Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
DB 37 AAPL 40

RESULT 26

C81840

hypothetical protein NMA1489 [imported] - Neisseria meningitidis (strain Z2491 serogroup

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: C81840

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: C81840

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-83 <PAR>

A:Cross-references: GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CAB84722.1; PID:g738013

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA1489

Query Match

100.0%; Score 19; DB 2; Length 83;

Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
DB 37 AAPL 40

RESULT 27

T35979

hypothetical protein SC9C7.28 - Streptomyces coelicolor (fragment)

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T35979

R:Seeger, K.U.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, January 1999

A:Reference number: Z21551

A:Accession: T35979
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-83 <SEED>
A:Cross-references: EMBL:AL035161; PIDN:CAA22740.1; GSPDB:GN00070; SCOEDB:SC9C7.28
C:Genetics:
A:Gene: SCOEDB:SC9C7.28

Query Match 100.0%; Score 19; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
DB 13 AAPL 16

RESULT 28

AI3550

hypothetical protein BMEII0330 [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C:Accession: AI3550

R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov

; Mazur, M.; Goltsman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Let

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit

A:Reference number: AD3252; PMID:11756688

A:Accession: AI3550

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-84 <KUR>

A:Cross-references: GB:AE008918; PIDN:AAL53572.1; PID:g17984483; GSPDB:GN00191

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEII0330

A:Map position: II

Query Match 100.0%; Score 19; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
DB 42 AAPL 45

RESULT 29

A39747

hypothetical protein (D-loop region) - minke whale mitochondrion

C:Species: mitochondrion Balaenoptera acutorostrata (minke whale, lesser rorqual)

C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 07-Dec-1999

C:Accession: A39747

R:Hoelzel, A.R.; Hancock, J.M.; Dover, G.A.

Mol. Biol. Evol. 8, 475-493, 1991

A:Title: Evolution of the cetacean mitochondrial D-loop region.

A:Reference number: A39747; MUID:92017215; PMID:1717809

A:Accession: A39747

A:Molecule type: DNA

A:Residues: 1-85 <HOE>

A:Cross-references: GB:M60408; NID:g336457; PIDN:AAA31673.1; PID:g336458

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC1

C:Keywords: mitochondrion

Query Match 100.0%; Score 19; DB 2; Length 85;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||

Db 52 AAPL 55

RESULT 30
C90666

probable DNA binding protein [imported] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-May-2002
C:Accession: C90666
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C90666
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <HAY>
A:Cross-references: GB:BA000007; PIDN:BAB33722.1; PID:g13359756; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs0299
C:Superfamily: Escherichia coli prophage cp4-57 regulatory protein alpha

Query Match 100.0%; Score 19; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 12 AAPL 15

RESULT 31
G85516

probable DNA binding protein [imported] - Escherichia coli (strain O157:H7, substrain ET
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
C:Accession: G85516
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: G85516
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <STO>
A:Cross-references: GB:AE005174; MID:g12513050; PIDN:AAG54595.1; GSPDB:GN00145; UWGP:Z03
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z0336
C:Superfamily: Escherichia coli prophage cp4-57 regulatory protein alpha

Query Match 100.0%; Score 19; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 12 AAPL 15

RESULT 32
PC4061

pyrrolo-quinoline-quinone protein III - Acinetobacter calcoaceticus (fragment)
C:Species: Acinetobacter calcoaceticus
C:Date: 15-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 08-Oct-1999
C:Accession: PC4061
R:Adachi, H.; Tsujimoto, M.
J. Biochem. 118, 555-561, 1995
A:Title: Cloning and expression of dipeptidase from Acinetobacter calcoaceticus ATCC 230
A:Reference number: JC4222; MUID:96115926; PMID:8690717

A:Accession: PC4061
A:Molecule type: DNA
A:Residues: 1-88 <ADA>
A:Cross-references: DDBJ:D50330; MID:g1088398; PIDN:BAA08866.1; PID:d1009499; PID:g10
C:Genetics:
A:Gene: pqiIII

Query Match 100.0%; Score 19; DB 2; Length 88;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 57 AAPL 60

RESULT 33
C36869

probable export protein fliQ - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002
C:Accession: C36869; B64959
R:Malakooti, J.; Ely, B.; Matsumura, P.
J. Bacteriol. 176, 189-197, 1994
A:Title: Molecular characterization, nucleotide sequence, and expression of the fliQ,
A:Reference number: A36869; MUID:94110225; PMID:8282695
A:Accession: C36869
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-89 <MAL>

A:Cross-references: GB:L22182; MID:g347241; PIDN:AC36860.1; PID:g347244
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64959
A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-89 <BLAT>
A:Cross-references: GB:AE000287; GB:U00096; MID:g1788257; PIDN:AACT5016.1; PID:g17882
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: fliQ
C:Function:
A:Description: required for assembly of the flagellum rivet; may be involved in expor
C:Superfamily: flagellar biosynthesis-specific protein
C:Keywords: flagellum; transmembrane protein
F:22-38/Domain: transmembrane #status predicted <TM1>
F:49-65/Domain: transmembrane #status predicted <TM2>

Query Match 100.0%; Score 19; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 21 AAPL 24

RESULT 34
S78699

probable export protein fliQ - Salmonella typhimurium
C:Species: Salmonella typhimurium
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: S78699
R:Ohnishi, K.; Fan, F.; Schoenhals, G.J.; Kihara, M.; Macnab, R.M.
J. Bacteriol. 179, 6092-6099, 1997
A:Title: The fliQ, fliP, and fliR proteins of Salmonella typhimurium: putative
A:Reference number: S78696; MUID:97464436; PMID:9324257
A:Accession: S78699
A:Molecule type: DNA

A;Residues: 1-89 <OHN>
A;Cross-references: EMBL:L49021; NID:g1066860; PIDN:AAB81320.1; PID:g1066864
A;Note: the sequence of residues 27-59 and the corresponding nucleic acid sequence are n
C;Genetics: fltQ
A;Gene: fltQ
C;Function:
A;Description: required for assembly of the flagellum rivet; may be involved in export c
C;Superfamily: flagellar biosynthesis-specific protein
C;Keywords: flagellum; transmembrane protein
F;22-38/Domain: transmembrane #status predicted <TM1>
F;49-65/Domain: transmembrane #status predicted <TM2>

Query Match 100.0%; Score 19; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
Db 21 AAPL 24

RESULT 35
AE0753
flagellar biosynthetic protein fltQ [imported] - Salmoneilla enterica subsp. enterica ser
C;Species: Salmoneilla enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmoneilla typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 17-May-2002
C;Accession: AE0753
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmoneilla enterica serov
A;Reference number: AB0502; PMID:11677608
A;Accession: AE0753
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05728.1; PID:g16503221; GSPDB:GN00176
C;Genetics:
A;Gene: STY2188
C;Superfamily: flagellar biosynthesis-specific protein

Query Match 100.0%; Score 19; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
Db 21 AAPL 24

RESULT 36
AI0221
flagellar biosynthetic protein fltQ [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 17-May-2002
C;Accession: AI0221
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tilball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AI0221
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC90637.1; PID:g15979843; GSPDB:GN00175
C;Genetics:
A;Gene: fltQ
C;Superfamily: flagellar biosynthesis-specific protein

Query Match 100.0%; Score 19; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
Db 21 AAPL 24

RESULT 37
H90964
probable export protein fltQ [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: H90964
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H90964
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA036111.1; PID:g13362156; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECS2688

Query Match 100.0%; Score 19; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
Db 21 AAPL 24

RESULT 38
H85812
flagellar biosynthesis [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: H85812
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
hler, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimantanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: AB5480; MUID:21074935; PMID:11206551
A;Accession: H85812
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <STO>
A;Cross-references: GB:AE005174; NID:g12516056; PIDN:AAG56964.1; GSPDB:GN00145; UWGP:
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: fltQ

Query Match 100.0%; Score 19; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
Db 21 AAPL 24

RESULT 39
T31980
hypothetical protein D2062.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T31980
 R:Tin-Wollam, A.; Wohlmann, P.
 submitted to the EMBL Data Library, July 1997
 A:Description: The sequence of C. elegans cosmid D2062.
 A:Reference number: 221107
 A:Accession: T31980
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-90 <TIN>
 A:Cross-references: EMBL:AF016664; PIDN:AAB66067.1; GSPDB:GN00020; CESP:D2062.6
 A:Experimental source: strain Bristol N2; clone D2062
 C:Genetics:
 A:Gene: CESP:D2062.6
 A:Map position: 2

Query Match 100.0%; Score 19; DB 2; Length 90;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AAPL 4
 ||||
 Db 44 AAPL 47

RESULT 40

H69198
 hypothetical protein MTH74 - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: H69198
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 ; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
 K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: H69198
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-91 <MTH>
 A:Cross-references: GB:AE000666; MID:g2621094; PIDN:AAB84578.1; PID:g2621108
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH74
 C:Superfamily: Methanobacterium hypothetical protein MTH80

Query Match 100.0%; Score 19; DB 1; Length 91;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AAPL 4
 ||||
 Db 48 AAPL 51

Search completed: December 6, 2002, 13:31:02
 Job time : 10.2222 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:20:29 ; Search time 5.55556 Seconds
(without alignments)
29.863 Million cell updates/sec

Title: US-10-033-526-3
Perfect score: 19
Sequence: 1 AAPL 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 19 | 100.0 | 30 | 1 Y161_TREPA | O83196 treponema p |
| 2 | 19 | 100.0 | 36 | 1 PSBY_ODOSI | P49543 odontella s |
| 3 | 19 | 100.0 | 53 | 1 Y871_TREPA | O83841 treponema p |
| 4 | 19 | 100.0 | 67 | 1 CXET_CONTE | P81755 conus texti |
| 5 | 19 | 100.0 | 75 | 1 HPIS_RHOFE | P80882 rhodofera |
| 6 | 19 | 100.0 | 76 | 1 CONG_CONGE | O9xyr5 conus geogr |
| 7 | 19 | 100.0 | 81 | 1 MCP1_HIRME | P81511 hirudo medi |
| 8 | 19 | 100.0 | 89 | 1 FLIO_ECOLI | P33134 escherichia |
| 9 | 19 | 100.0 | 89 | 1 FLIO_SALTY | P54701 salmonella |
| 10 | 19 | 100.0 | 92 | 1 SI06_CHICK | O98953 gallus gall |
| 11 | 19 | 100.0 | 96 | 1 NULM_PETMA | O35541 petromyzon |
| 12 | 19 | 100.0 | 96 | 1 YL50_DEIRA | O9rsh4 deinococcus |
| 13 | 19 | 100.0 | 107 | 1 HITA_NEIGO | O07817 neisseria 9 |
| 14 | 19 | 100.0 | 107 | 1 SS13_STRCO | P29608 streptomyc |
| 15 | 19 | 100.0 | 109 | 1 CYPC_STRHA | O05368 streptomyc |
| 16 | 19 | 100.0 | 110 | 1 RLAI_ALTAL | P49148 alternaria |
| 17 | 19 | 100.0 | 117 | 1 NLTP_BETVU | O43748 beta vulgar |
| 18 | 19 | 100.0 | 117 | 1 YBAE_ECOLI | P09161 escherichia |
| 19 | 19 | 100.0 | 119 | 1 LAMR_PETMA | P33577 petromyzon |
| 20 | 19 | 100.0 | 121 | 1 FLIO_ECOLI | P22586 escherichia |
| 21 | 19 | 100.0 | 121 | 1 RL24_PYRAB | O9vlu7 pyrococcus |
| 22 | 19 | 100.0 | 125 | 1 RL7_RICPR | O9ze21 rickettsia |
| 23 | 19 | 100.0 | 129 | 1 NIFA_AZOC | P56269 azotobacter |
| 24 | 19 | 100.0 | 130 | 1 Y700_TREPA | O83698 treponema p |
| 25 | 19 | 100.0 | 132 | 1 YD69_SCHPO | O10321 schizosacch |
| 26 | 19 | 100.0 | 134 | 1 CU3A_TENMO | P80683 tenebrio mo |
| 27 | 19 | 100.0 | 139 | 1 LAMP_PETMA | P33575 petromyzon |
| 28 | 19 | 100.0 | 141 | 1 EXBD_ECOLI | P18784 escherichia |
| 29 | 19 | 100.0 | 141 | 1 ZRAP_ECOLI | P32682 escherichia |
| 30 | 19 | 100.0 | 142 | 1 EXBD_PSEPU | O05606 pseudomonas |
| 31 | 19 | 100.0 | 144 | 1 SREL_PIG | O97676 sus scrofa |
| 32 | 19 | 100.0 | 144 | 1 SSI_STRAO | P01006 streptomyc |
| 33 | 19 | 100.0 | 144 | 1 SSI_STRCO | P28591 streptomyc |

| | | | | | |
|----|----|-------|-----|--------------|--------------------|
| 34 | 19 | 100.0 | 147 | 1 YE86_MYCPN | P75299 mycoplasma |
| 35 | 19 | 100.0 | 149 | 1 RL9_MYCPN | P75540 mycoplasma |
| 36 | 19 | 100.0 | 150 | 1 EXON_PRYN3 | P30660 pseudorabie |
| 37 | 19 | 100.0 | 150 | 1 RL9_MYCGE | P47339 mycoplasma |
| 38 | 19 | 100.0 | 150 | 1 TTHY_CHICK | P27731 gallus gall |
| 39 | 19 | 100.0 | 150 | 1 TTHY_CROPO | O55245 crocodylus |
| 40 | 19 | 100.0 | 150 | 1 TTHY_TILRU | P30623 tiliqua rug |
| 41 | 19 | 100.0 | 150 | 1 YE42_MYCPN | P75336 mycoplasma |
| 42 | 19 | 100.0 | 151 | 1 YPR3_ECOLI | P10027 escherichia |
| 43 | 19 | 100.0 | 153 | 1 NODC_BRASP | P04677 bradyrhizob |
| 44 | 19 | 100.0 | 155 | 1 FLIL_SALTY | P26417 salmonella |
| 45 | 19 | 100.0 | 156 | 1 CU55_ARADI | P80518 araneus dia |

ALIGNMENTS

```
RESULT 1
Y161_TREPA
ID Y161_TREPA STANDARD; PRT; 30 AA.
AC O83196;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0161.
GN TP0161.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artlach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RT Science 281:375-388(1998).
RL -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001200; AAC65153.1; -.
DR TIGR; TP0161; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 30 AA; 3259 MW; 4B9A413A5498D15F CRC64;

Query Match 100.0%; Score 19; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 6 AAPL 9

RESULT 2
PSBY_ODOSI
ID PSBY_ODOSI STANDARD; PRT; 36 AA.
AC P49543;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
```

```
DE Photosystem II protein Y.
GN (PSBY-A OR PSBY-1 OR YCF32.1) AND (PSBY-B OR PSBY-2 OR YCF32.2).
OS Odontella sinensis (Marine centric diatom).
OG Chloroplast.
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Biddulphiophycidae; Eupodiscales; Eupodiscaceae; Odontella.
OX NCBI_TaxID=2839;
RN [1]
RP SEQUENCE FROM N.A.
RA Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RT Odontella sinensis.";
RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC -1- FUNCTION: MANGANESE-BINDING POLYPEPTIDE WITH L-ARGININE
CC METABOLIZING ENZYME ACTIVITY. COMPONENT OF THE CORE OF PHOTOSYSTEM
CC II (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -1- SIMILARITY: BELONGS TO THE PSBY FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; Z67753; CAA91612.1; -.
CC DR EMBL; Z67753; CAA91666.1; -.
CC KW Photosystem II; Chloroplast; Transmembrane; Thylakoid.
CC FT DOMAIN 1 6 LUMENAL (POTENTIAL).
CC FT TRANSMEM 7 23 POTENTIAL.
CC FT DOMAIN 24 36 STROMAL (POTENTIAL).
CC SQ SEQUENCE 36 AA; 3921 MW; 22989D3E7DAE2B63 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AAPL 4
Db 9 AAPL 12

RESULT 3
Y871_TREPA STANDARD; PRT; 53 AA.
AC 083841;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0871.
GN TP0871.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artlach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AE001257; AAC65846.1; -.
CC DR TIGR; TP0871; -.
CC KW Hypothetical protein; Complete proteome.
CC SQ SEQUENCE 53 AA; 5998 MW; F238BDAFD5C6D6D7 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AAPL 4
Db 42 AAPL 45

RESULT 4
CXET_CONTE STANDARD; PRT; 67 AA.
AC P81755; Q9U627;
DT 30-MAY-2000 (Rel. 39, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epsilon-conotoxin Txix precursor (Conotoxin tx5a).
OS Conus textile (Cloth-of-gold cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 51-63, AND MASS SPECTROMETRY.
RC TISSUE=Venom duct, and Venom;
RX MEDLINE=99452958; PubMed=10521453;
RA Walker C.S., Steel D., Jacobsen R.B., Lirazan M.B., Cruz L.J.,
RA Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,
RA Bandyopadhyay P., Craig A.G., Olivera B.M.;
RA J. Biol. Chem. 274:36030-36030(1999).
RL [2]
RN [2]
RP ERRATUM.
RA Walker C.S., Steel D., Jacobsen R.B., Lirazan M.B., Cruz L.J.,
RA Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,
RA Bandyopadhyay P., Craig A.G., Olivera B.M.;
RA J. Biol. Chem. 274:36030-36030(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom duct;
RX MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for evolving hypervariability: the case of conopeptides.";
RL Mol. Biol. Evol. 18:120-131(2001).
RN [4]
RP SEQUENCE OF 51-63, CARBOHYDRATE-LINKAGE SITE THR-10, AND STRUCTURE BY
RP NMR.
RC TISSUE=Venom;
RX MEDLINE=99254114; PubMed=10318957;
RA Rigby A.C., Lucas-Meunier E., Kalume D.E., Czerwiec E., Hambe B.,
RA Dahlqvist I., Fossier P., Baux G., Roepstorff P., Baleja J.D.,
RA Furlie B.C., Furlie B., Stenflo J.P.;
RT "A conotoxin from Conus textile with unusual posttranslational
RT modifications reduces presynaptic Ca2+ influx.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:5758-5763(1999).
RN [5]
RP SEQUENCE OF 51-63, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX PubMed=10679974;
RA Kalume D.E., Stenflo J.P., Czerwiec E., Hambe B., Furlie B.C.,
RA Furlie B., Roepstorff P.;
RT "Structure determination of two conotoxins from Conus textile by a
```

```

RT combination of matrix-assisted laser desorption/ionization
RT time-of-flight and electrospray ionization mass spectrometry and
RT biochemical methods."
RL J. Mass Spectrom. 35:145-156(2000).
CC -1- FUNCTION: Acts at presynaptic membranes, blocking the calcium
CC channels. Causes hyperactivity upon intracranially injection into
CC mice. Causes dorsal fins drooping in fish.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- PTM: O-glycan consists of the disaccharide Gal-GalNAc.
CC -1- MASS SPECTROMETRY: MW=1929.4; METHOD=Electrospray.
CC -1- MASS SPECTROMETRY: MW=1929.4; METHOD=MALDI.
CC -1- SIMILARITY: BELONGS TO THE T-SUPERFAMILY OF CONOTOXINS. EPSILON-
CC TYPE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF167167; AAF03687.1; -.
DR EMBL; AF214958; AAG60386.1; -.
DR PDB; 1WCT; 08-JUN-99.
KW Presynaptic neurotoxin; Neurotoxin; Toxin; Calcium channel inhibitor;
KW Vitamin K; Signal; Gamma-carboxyglutamic acid; Glycoprotein;
KW Hydroxylation; Bromination; 3D-structure.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 50
FT PEPTIDE 51 63 EPSILON-CONOTOXIN TXIX.
FT PROPEP 64 67
FT DISULFID 52 58
FT DISULFID 53 59
FT MOD_RES 51 51
FT MOD_RES 54 54 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 57 57 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 63 63 BROMINATION.
FT CARBOHYD 60 60 O-LINKED (GALNAC. . .).
SQ SEQUENCE 67 AA; 7587 MW; 7270505504D6B3D CRC64;

Query Match 100.0%; Score 19; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 61 AAPL 64

RESULT 5
HPIS_RHOFE
ID HPIS_RHOFE STANDARD; PRT; 75 AA.
AC P80882;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High potential iron-sulfur protein (HiPIP).
GN HiP.
OS Rhodoferrax fermentans.
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC Rhodoferrax.
OX NCBI_TaxID=28066; -
RN [1]
RP SEQUENCE.
RC STRAIN=JMC 7819;
RX MEDLINE=97234565; PubMed=9119002;
RA van Driessche G., Ciurli S., Hochkoeppler A., van Beeumen J.J.;
RT "The primary structure of Rhodoferrax fermentans high-potential iron-
RT sulfur protein, an electron donor to the photosynthetic reaction
RT center.";
RL Eur. J. Biochem. 244:371-377(1997).

```

```

CC -1- FUNCTION: SPECIFIC CLASS OF HIGH-REDOX-POTENTIAL 4FE-4S
CC FERREDOXINS. FUNCTIONS IN ANAEROBIC ELECTRON TRANSPORT IN MOST
CC PURPLE AND IN SOME OTHER PHOTOSYNTHETIC BACTERIA AND IN AT LEAST
CC ONE GENUS (PARACOCUS) OF HALOPHILIC, DENITRIFYING BACTERIA.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- MASS SPECTROMETRY: MW=7849.64; METHOD=Electrospray.
CC -1- SIMILARITY: BELONGS TO THE HIGH POTENTIAL IRON-SULFUR PROTEIN
CC (HiPIP) FAMILY.
CC HSSP: P00260; ICKU.
DR InterPro; IPR000170; Hipot_Ironsulf.
DR Pfam; PF01355; HiPIP; 1.
DR PROSITE; PS00596; HiPIP; 1.
KW Electron transport; Iron-sulfur; 4Fe-4S.
FT METAL 38 38 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 41 41 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 54 54 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 68 68 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 75 AA; 7849 MW; 9FE6C86E62A72D3 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 75;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 1 AAPL 4

RESULT 6
CONG_CONGE
ID CONG_CONGE STANDARD; PRT; 76 AA.
AC Q9XYR5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Contulakin-G precursor.
OS Conus geographus (Geography cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6491;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 51-66, AND CHARACTERIZATION.
RC TISSUE=Venom;
RX MEDLINE=99253920; PubMed=10318778;
RA Craig A.G., Norberg T., Griffin D., Hoeger C., Akhtar M., Schmidt K.,
RA Low W., Dykert J., Richelson E., Navarro V., Mazella J., Watkins M.,
RA Hillyard D.R., Imperial J., Cruz L.J., Olivera B.M.;
RT "Contulakin-G, an O-glycosylated invertebrate neurotensin.";
RL J. Biol. Chem. 274:13752-13759(1999).
CC -1- FUNCTION: Acts as an agonist of neurotensin receptors.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- PTM: O-glycosylated. The glycosylation seems to enhance
CC the affinity to the neurotensin receptors.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF121108; AAD30031.1; -.
DR GLCOSuitedB; Q9XYR5; -.
KW Toxin; Glycoprotein; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 48
FT PEPTIDE 51 66 CONTULAKIN-G.
FT PROPEP 67 76
FT MOD_RES 51 51 PYRROLIDONE CARBOXYLIC ACID.
FT CARBOHYD 60 60 O-LINKED (GALNAC. . .).
FT /FTId=CAR_000164.

```

SO SEQUENCE 76 AA; 8261 MW; D8094EB30C6AAFD4 CRC64;
Query Match 100.0%; Score 19; DB 1; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPL 4
Db 16 AAPL 19
RESULT 7
MCPI_HIRME STANDARD; PRT; 81 AA.
AC P81511;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Metallocarboxypeptidase inhibitor precursor (LCI).
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 16-81.
RX MEDLINE=99047680; PubMed=9830043;
RA Reverter D., Vendrell J., Canals F., Horstmann J., Aviles F.X.,
RA Fritz H., Sommerhoff C.P.;
RT "A carboxypeptidase inhibitor from the medical leech Hirudo
RT medicinalis. Isolation, sequence analysis, cDNA cloning, recombinant
RT expression, and characterization.";
RL J. Biol. Chem. 273:32927-32933(1998).
CC -1- FUNCTION: TIGHTLY BINDING, COMPETITIVE INHIBITOR OF DIFFERENT
CC TYPES OF PANCREATIC-LIKE CARBOXYPEPTIDASES.
CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
CC BONDS.
CC -1- MASS SPECTROMETRY: MW=7326; METHOD=MALDI; RANGE=16-81.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ010948; CAA09422.1; .
KW Metalloenzyme inhibitor; signal.
FT SIGNAL 1 15
FT CHAIN 16 81 METALLOCARBOXYPEPTIDASE INHIBITOR.
FT ACT_SITE 80 80 INTERACTION WITH CARBOXYPEPTIDASE (BY
FT SIMILAR 75 80 SIMILARITY).
FT TO THE C-TERMINAL OF SOLANACEA
FT CARBOXYPEPTIDASE INHIBITORS.
SQ SEQUENCE 81 AA; 9068 MW; 7D7E175E6933922A CRC64;
Query Match 100.0%; Score 19; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPL 4
Db 39 AAPL 42
RESULT 8
FLIQ_ECOLI STANDARD; PRT; 89 AA.
ID FLIQ_ECOLI
AC P33134;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar biosynthetic protein flIQ.

GN FLIQ OR FLAQ OR B1949 OR Z3039 OR ECS2688.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=K12;
RC STRAIN=K12;
RX MEDLINE=94110225; PubMed=8282695;
RA Malakooti J., Ely B., Matsumura P.;
RT "Molecular characterization, nucleotide sequence, and expression of
RT the flIQ, flIP, flIQ, and flIR genes of Escherichia coli.";
RL J. Bacteriol. 176:189-197(1994).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RC STRAIN=K12;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=K12;
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Maki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampei G., Seki Y., Sivasubaram S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [5]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / RIMD 0509952;
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -1- FUNCTION: REQUIRED FOR THE ASSEMBLY OF THE RIVET AT THE EARLIEST
CC STAGE OF FLAGELLAR BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -1- SIMILARITY: BELONGS TO THE FLIQ/MOPD/SPAQ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----


```

CC -----
DR EMBL; L22182; AAC36860.1; -.
DR EMBL; AE000287; AAC75016.1; -.
DR EMBL; D90834; BAA15774.1; -.
DR EMBL; D90835; BAA15780.1; -.
DR EMBL; AE005417; AAG36964.1; -.
DR EMBL; AP002559; BAB36111.1; -.
DR EcoGene; EG11976; fltQ.
DR InterPro; IPR002191; Bac_export_3.
DR Pfam; PF01313; Bac_export_3; 1.
DR PRINTS; PR00952; TYPE3IMQPROT.
KW Flagella; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 16 40 POTENTIAL.
FT TRANSMEM 55 75 POTENTIAL.
SQ SEQUENCE 89 AA; 9632 MW; 8ED2B04C69BCD5AD CRC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 1; Length 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
   1111
Db 21 AAPL 24

RESULT 9
FLIQ_SALTY STANDARD; PRT; 89 AA.
AC P54701;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flagellar biosynthetic protein fltQ.
GN fltQ OR FLAQ OR STM1980 OR STY2188.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2;
RX MEDLINE=97464436; PubMed=9324257;
RA Ohnishi K., Fan F., Schoenhals G.J., Kihara M., Macnab R.M.;
RT "The fltQ, fltP, fltQ, and fltR proteins of Salmonella typhimurium:
RT putative components for flagellar assembly.";
RL J. Bacteriol. 179:6092-6099(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford B.G., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
```

```

RL Nature 413:848-852(2001).
CC -!- FUNCTION: REQUIRED FOR THE ASSEMBLY OF THE RIVET AT THE EARLIEST
CC STAGE OF FLAGELLAR BIOSYNTHESIS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- SIMILARITY: BELONGS TO THE FLIQ/MOPD/SPAQ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L49021; AAB81320.1; -.
DR EMBL; AE008787; AAL20892.1; -.
DR EMBL; AL627272; CAD05728.1; -.
DR STGene; SG10578; fltQ.
DR InterPro; IPR002191; Bac_export_3.
DR Pfam; PF01313; Bac_export_3; 1.
DR PRINTS; PR00952; TYPE3IMQPROT.
KW Flagella; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 18 38 POTENTIAL.
FT TRANSMEM 55 75 POTENTIAL.
SQ SEQUENCE 89 AA; 9604 MW; 95DFF987E6BCCAB6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 1; Length 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
   1111
Db 21 AAPL 24

RESULT 10
S106_CHICK STANDARD; PRT; 92 AA.
ID S106_CHICK
AC Q98953;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcyclin.
GN S100A6.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gizzard;
RX MEDLINE=98262396; PubMed=9599662;
RA Allen B.G., Andrea J.E., Sutherland C., Schonekess B.O., Walsh M.P.;
RT "Molecular cloning of chicken calcyclin (S100A6) and identification
RT of putative isoforms.";
RL Biochem. Cell Biol. 75:733-738(1997).
RL -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC -!- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U76365; AAB18788.1; -.
DR HSSP; P30801; 1A03.
DR InterPro; IPR001751; GABP_S100.
DR InterPro; IPR002048; EF-hand.
```



```
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CABP_S100; 1.
DR ProSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Mitogen; Cell cycle; Calcium-binding.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ SEQUENCE 92 AA; 10276 MW; 56B3D65BAA7BF7A4 CRC64;
```

```
Query Match 100.0%; Score 19; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAPL 4
    ||||
Db 2 AAPL 5
```

```
RESULT 11
NULM_PETMA STANDARD; PRT; 96 AA.
ID NULM_PETMA
```

```
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L OR NAD4L.
OS Petromyzon marinus (Sea lamprey).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
```

```
RP [1]
RX MEDLINE=95229067; PubMed=7713438;
RA Lee W.J., Kocher T.D.;
RT "Complete sequence of a sea lamprey (Petromyzon marinus)
RT mitochondrial genome: early establishment of the vertebrate genome
RT organization."
RL Genetics 139:873-887(1995).
```

```
CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; U11880; AAB08746.1; -
DR InterPro; IPR001133; Oxidored_4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 96 AA; 10392 MW; 4E042827931C8BC8 CRC64;
```

```
Query Match 100.0%; Score 19; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAPL 4
    ||||
Db 56 AAPL 59
```

```
RESULT 12
YL50_DEIRA STANDARD; PRT; 96 AA.
ID YL50_DEIRA
AC Q9RSH4;
DT 16-OCT-2001 (Rel. 40, Created)
```

```
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein DR2150.
GN DR2150.
```

```
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococi; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
```

```
CC -I- SIMILARITY: BELONGS TO THE UPE0161 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; AE002049; AAF11698.1; -
DR TIGR; DR2150; -
DR InterPro; IPR002696; DUF37.
DR Pfam; PF01809; DUF37; 1.
DR ProDom; PD004225; DUF37; 1.
DR TIGRFAMS; TIGR00278; DUF37; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 96 AA; 10628 MW; EA1201670EE9D1E3 CRC64;
```

```
Query Match 100.0%; Score 19; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAPL 4
    ||||
Db 3 AAPL 6
```

```
RESULT 13
HITA_NEIGO STANDARD; PRT; 107 AA.
ID HITA_NEIGO
AC 007817;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HITA protein.
GN HITA.
```

```
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F62;
RA McGee D.J., Srivastava R., Rest R.F.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: BELONGS TO THE HIT FAMILY.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF003550; AAB61288.1; -.
DR HSSP; P49773; IKPC.
DR InterPro; IPR001310; HIT.
DR Pfam; PF01230; HIT; 1.
DR PRINTS; PR00332; HISTRIAD.
DR PROSITE; PS00892; HIT; 1.
SQ SEQUENCE 107 AA; 11612 MW; 1315960D0912F86B CRC64;

Query Match 100.0%; Score 19; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
   ||||
Db 34 AAPL 37

RESULT 14
SSI3_STRCO STANDARD; PRT; 107 AA.
ID SSI3_STRCO
AC P29608; O9R5B8;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Subtilisin inhibitor-like protein-3 (SIL-3) (SIL3).
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE.
RC STRAIN=KCC S006;
RX MEDLINE=94192683; PubMed=8143745;
RA Taguchi S., Kojima S., Terabe M., Miura K.-I., Momose H.;
RT "Comparative studies on the primary structures and inhibitory
RT properties of subtilisin-trypsin inhibitors from Streptomyces.";
RL Eur. J. Biochem. 220:911-918(1994).
RN [2]
RP SEQUENCE OF 1-35.
RC STRAIN=KCC S006;
RX MEDLINE=93146392; PubMed=1490613;
RA Taguchi S., Kojima S., Kumagai I., Ogawara H., Miura K.-I., Momose H.;
RT "Isolation and partial characterization of SSI-like protease
RT inhibitors from Streptomyces.";
RL FEBS Microbiol. Lett. 78:293-297(1992).
RN [3]
RP SEQUENCE OF 1-34.
RC STRAIN=KCC S006;
RX MEDLINE=93222542; PubMed=7763545;
RA Taguchi S., Kikuchi H., Kojima S., Kumagai I., Nakase T., Miura K.-I.,
RA Momose H.;
RT "High frequency of SSI-like protease inhibitors among Streptomyces.";
RL Biosci. Biotechnol. Biochem. 57:522-524(1993).
CC -1- FUNCTION: INHIBITOR OF SUBTILISIN BPN' AND TRYPSIN.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: secreted.
CC -1- SIMILARITY: BELONGS TO THE SERINE PROTEASE INHIBITORS SSI FAMILY.
DR PIR; S42572; S42572.
DR HSSP; P01006; 2SIC.
DR InterPro; IPR000691; Strep_subt_inhib.
DR Pfam; PF00720; SSI; 1.
DR PRINTS; PR00294; SSBTLNINHBT.
DR PRODOM; PD004028; Strep_subt_inhib; 1.
DR PROSITE; PS00999; SSI; 1.
KW Serine protease inhibitor.
FT DISULFID 29 44 BY SIMILARITY.
FT DISULFID 65 95 BY SIMILARITY.
FT ACT_SITE 67 68 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 107 AA; 10924 MW; 71C84BCDF44355DE CRC64;
```

```
Query Match 100.0%; Score 19; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
   ||||
Db 19 AAPL 22

RESULT 15
CYPC_STRHA STANDARD; PRT; 109 AA.
ID CYPC_STRHA
AC O05368;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative polyketide cyclase (Fragment).
GN SCH4.
OS Streptomyces halstedii.
OC Bacteria; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1944;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2381;
RX MEDLINE=93345807; PubMed=8344517;
RA Blanco G., Brian P., Pereda A., Mendez C., Salas J.A., Chater K.F.;
RT "Hybridization and DNA sequence analyses suggest an early
RT evolutionary divergence of related biosynthetic gene sets encoding
RT polyketide antibiotics and spore pigments in Streptomyces spp.";
RL Gene 130:107-116(1993).
CC -1- FUNCTION: INVOLVED IN DEVELOPMENTALLY REGULATED SYNTHESIS OF A
CC COMPOUND BIOSYNTHETICALLY RELATED TO POLYKETIDE ANTIBIOTICS,
CC WHICH IS ESSENTIAL FOR SPORE COLOR IN STREPTOMYCES HALSTEDII.
CC -1- SIMILARITY: TO POLYKETIDE CYCLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L05390; AAA02836.2; -.
DR PIR; PN0640; PN0640.
FT NON_TER 109 109
SQ SEQUENCE 109 AA; 12862 MW; 2DC56818A4319950 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
   ||||
Db 12 AAPL 15

RESULT 16
RLAL_ALTAL STANDARD; PRT; 110 AA.
ID RLAL_ALTAL
AC P49148;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 60S acidic ribosomal protein P1 (Allergen Alt a 12) (Alt a XII).
GN ALTA12.
OS Alternaria alternata.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC Pleosporales; Pleosporaceae; mitosporic Pleosporaceae; Alternaria.
OX NCBI_TaxID=5599;
RN [1]
```

RP SEQUENCE FROM N.A.
RC STRAIN-08-0203-Berlin;
RX MEDLINE=95206305; PubMed=7898496;
RA Achatz G., Oberkofler H., Lechenauer E., Simon B., Unger A.,
RA Kandler D., Ebner C., Prillinger H., Kraft D., Breitenbach M.;
RT "Molecular cloning of major and minor allergens of Alternaria
alternata and Cladosporium herbarum.";
RL Mol. Immunol. 32:213-227(1995).
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
CC PROTEIN SYNTHESIS.
CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
CC SUBUNIT.
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X84216; CAA58998.1; -
DR InterPro; IPR001813; 60S_ribosomal.
DR Pfam; PF00428; 60S_ribosomal; 1.
KW Ribosomal protein; Phosphorylation; Allergen.
SQ SEQUENCE 110 AA; 11736 MW; F06A232B67C3259F CRC64;

Query Match 100.0%; Score 19; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
| | | |
Db 69 AAPL 72

RESULT 17
NLTP_BETVU
ID NLTP_BETVU STANDARD; PRT; 117 AA.
AC Q43748;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nonspecific lipid-transfer protein precursor (LTP).
GN IMF1' OR IMFA.
OS Beta vulgaris (Sugar beet).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
OX NCBI_TaxID=3555;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Monova; TISSUE=Leaf;
RX MEDLINE=96382423; PubMed=8790287;
RA Nielsen K.K., Nielsen J.E., Madrid S.M., Mikkelsen J.D.;
RT "New antifungal proteins from sugar beet (Beta vulgaris L.) showing
RT homology to non-specific lipid transfer proteins.";
RL Plant Mol. Biol. 31:539-552(1996).
CC -1- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEINS TRANSFER
CC PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY
CC A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING
CC EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES. ALSO HAS FUNGICIDE
CC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC -----
DR EMBL; X92748; CAA63407.1; -
DR HSP; P23096; 1R2L.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR000528; Plant_LTP.
DR InterPro; IPR001768; Try/amyL_inhbr.
DR Pfam; PF00234; tryp_alpha_amy1; 1.
DR PRINTS; PR00382; LIPIDTRANSFER.
DR SMART; SM00499; AAI; 1.
DR PROSITE; PS00597; PLANT_LTP; 1.
KW Plant defense; Fungicide; Lipid-binding; Transport; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 117 NONSPECIFIC LIPID-TRANSFER PROTEIN.
FT DISULFID 29 76 POTENTIAL.
FT DISULFID 39 53 POTENTIAL.
FT DISULFID 54 99 POTENTIAL.
FT DISULFID 74 113 POTENTIAL.
SQ SEQUENCE 117 AA; 11645 MW; C1A371D2DE68C636 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
| | | |
Db 20 AAPL 23

RESULT 18
YBAA_ECOLI
ID YBAA_ECOLI STANDARD; PRT; 117 AA.
AC P09161;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ybaA.
GN YBAA OR B0456 OR Z0568 OR ECS0509.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85033714; PubMed=6208370;
RA Hsu L.M., Zagorski J., Fournier M.J.;
RT "Cloning and sequence analysis of the Escherichia coli 4.5 S RNA
RT gene.";
RL J. Mol. Biol. 178:509-531(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
RA Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,

```
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01074; CA25537.1; -.
DR EMBL; AE000152; AAC73558.1; -.
DR EMBL; U82664; AAB40211.1; -.
DR EMBL; AE005224; AAG54805.1; -.
DR EMBL; AP002551; BAB33932.1; -.
DR PIR; S07261; S07261.
DR EcoGene; EGI1099; ybaA.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 117 AA; 13318 MW; 5E320D62A2A93FF9 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 26 AAPL 29

RESULT 19
LAMR_PETMA
ID LAMR_PETMA STANDARD; PRT; 119 AA.
AC P33577;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Lamprin 1.8-10 precursor (Cartilage matrix protein).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-44.
RC TISSUE=Cartilage;
RX MEDLINE=93123269; PubMed=7678258;
RA Robson P., Wright G.M., Sitarz E., Maiti A., Rawat M., Youson J.H.,
RA Keeley F.W.;
RT "Characterization of lamprin, an unusual matrix protein from lamprey
RT cartilage. Implications for evolution, structure, and assembly of
RT elastin and other fibrillar proteins.";
RL J. Biol. Chem. 268:1440-1447(1993).
CC -----
CC -1- FUNCTION: SELF-AGGREGATING PROTEIN THAT MAKES PART OF THE SOLUBLE
CC FORM OF LAMPRIIN.
CC -1- SUBUNIT: THE POLYMERIC LAMPRIIN CHAINS SELF-AGGREGATE TO FORM
CC FIBERS AND HAVE SECONDARY STRUCTURES PARTICULARLY RICH IN BETA-
CC SHEETS AND IN BETA-TURNS.
CC -1- SUBCELLULAR LOCATION: Extracellular matrix.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L05926; AAA49270.1; -.
DR PIR; B45051; B45051.
KW Cartilage; Repeat; Connective tissue; Signal.
FT SIGNAL 1 19
FT CHAIN 20 119 LAMPRIIN 1.8-10.
FT DOMAIN 41 90 7 X 5 AA APPROXIMATE REPEATS.
FT REPEAT 41 45 1.
FT REPEAT 46 50 2.
FT REPEAT 51 55 3.
FT REPEAT 56 60 4.
FT REPEAT 61 65 5.
FT REPEAT 66 70 6.
FT REPEAT 86 90 7.
SQ SEQUENCE 119 AA; 11329 MW; 4EB476FACB7591B8 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 109 AAPL 112

RESULT 20
FLIO_ECOLI
ID FLIO_ECOLI STANDARD; PRT; 121 AA.
AC P22586; P33132;
DT 01-AUG-1991 (Rel. 19, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar protein flio.
GN FLIO OR FLBD OR FLAP OR B1947.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=94110225; PubMed=8282695;
RA Malakooti J., Ely B., Matsumura P.;
RT "Molecular characterization, nucleotide sequence, and expression of
RT the flio, flip, fliq, and flir genes of Escherichia coli.";
RL J. Bacteriol. 176:189-197(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampei G., Seki Y., Sivasubdaram S., Tagami H.,
RA Takeda J., Takenoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
```


RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE FLIO/MOPB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L22182; AAC36858.1; ALT_INIT.
DR EMBL; AE000287; AAC75014.1; ALT_INIT.
DR EMBL; D90834; BAA15772.1; ALT_INIT.
DR EMBL; D90835; BAA15778.1; -.
DR EcoGene; EG11224; fllo.
KW Chemotaxis; Flagella; Flagellar rotation; Transmembrane;
KM Complete proteome.
FT TRANSMEM 19 39 POTENTIAL.
FT CONFLICT 15 15 A -> R (IN REF. 1).
FT CONFLICT 18 18 L -> V (IN REF. 1).
SQ SEQUENCE 121 AA; 12670 MW; 0D4B6598368E60F5 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 121;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 15 AAPL 18

RESULT 21
RL24_PYRAB STANDARD; PRT; 121 AA.
ID RL24_PYRAB
AC Q9V1U7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L24P.
GN RPL24P OR PAB2128.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GE5 / Orsay;
RA Hellig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RT structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ248284; CAB49252.1; -
DR InterPro; IPR000302; KOW_motif.
DR Pfam; PF00467; KOW; 1.
DR ProDom; PD001677; Ribosomal_L24; 1.
DR TIGRFAMS; TIGR01080; rplX_A_E; 1.
DR PROSITE; PS01108; RIBOSOMAL_L24; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 121 AA; 14387 MW; 271D114AB3F21E67 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 121;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 28 AAPL 31

RESULT 22
RL7_RICPR STANDARD; PRT; 125 AA.
ID RL7_RICPR
AC Q9ZE21;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L7/L12.
GN RPL7 OR RP139.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sichertitz-Ponten T., Almark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."
RL Nature 396:133-140(1998).
CC -1- FUNCTION: SEEMS TO BE THE BINDING SITE FOR SEVERAL OF THE FACTORS
CC INVOLVED IN PROTEIN SYNTHESIS AND APPEARS TO BE ESSENTIAL FOR
CC ACCURATE TRANSLATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ235270; CAA14607.1; -
DR HSSP; P02392; ICTF.
DR InterPro; IPR000206; Ribosomal_L12.
DR Pfam; PF00542; Ribosomal_L12; 1.
DR ProDom; PD001326; Ribosomal_L12; 1.
DR TIGRFAMS; TIGR00855; L12; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 125 AA; 13213 MW; 431EDA82F25DB867 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 46 AAPL 49

RESULT 23
NIFA_AZOCH STANDARD; PRT; 129 AA.
ID NIFA_AZOCH
AC P56269;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Nif-specific regulatory protein (Fragment).
GN NIFA.
OS Azotobacter chroococcum mcd 1.


```
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Azotobacter.
OX NCBI_TaxID=355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89036145; PubMed=3053983;
RA Evans D., Jones R., Woodley P., Robson R.;
RT "Further analysis of nitrogen fixation (nif) genes in Azotobacter
RT chroococcum: identification and expression in Klebsiella pneumoniae
RT of nifs, nifV, nifM, and nifB genes and localization of nife/N-,
RT nifu-, nifa- and fixABC-like genes.";
RL J. Gen. Microbiol. 134:931-942(1988).
CC -i- FUNCTION: NIFA, A TRANSCRIPTIONAL ACTIVATOR, IS REQUIRED FOR
CC ACTIVATION OF MOST NIF OPERONS, WHICH ARE DIRECTLY INVOLVED IN
CC NITROGEN FIXATION. NIFA INTERACTS WITH SIGMA-54 (BY SIMILARITY).
CC -i- SIMILARITY: THE CENTRAL REGION CONTAINS A SIGMA-54 FACTOR
CC INTERACTION ATP-BINDING DOMAIN.
DR Interpro; IPR002197; HTH_FIS.
DR Interpro; IPR002078; Sig54_interact.
DR Pfam; PF02954; HTH_8; 1.
DR TIGRFAMS; TIGR01199; HTH_fis; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; PARTIAL.
DR PROSITE; PS00676; SIGMA54_INTERACT_2; PARTIAL.
DR PROSITE; PS00688; SIGMA54_INTERACT_3; PARTIAL.
DR PROSITE; PS50045; SIGMA54_INTERACT_4; 1.
KW Nitrogen fixation; transcription regulation; Activator;
KW ATP-binding; DNA-binding.
FT NON_TER 1 1
FT DOMAIN <1 46 SIGMA-54 FACTOR INTERACTION.
FT DNA_BIND 101 129 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 129 AA; 14425 MW; C30D30B18DF196D2 CRC64;
```

Query Match 100.0%; Score 19; DB 1; Length 129;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 71 AAPL 74

RESULT 24
Y700_TREPA STANDARD; PRT; 130 AA.
ID Y700_TREPA
AC O83698;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0700.
GN TP0700.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artlach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; AE001243; AAC65669.1; -.  
DR TIGR; TP0700; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 130 AA; 14204 MW; 09892AFCE31E4868 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 130;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



OY 1 AAPL 4  
Db 5 AAPL 8



RESULT 25  
YD69_SCHPO STANDARD; PRT; 132 AA.  
ID YD69_SCHPO  
AC Q10321;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Very hypothetical protein C17G8.09 in chromosome I.  
GN SPAC17G8.09.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volkart G., Aert R., Robben J., Grymptiez B.,  
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleaux V., Mottier S.,  
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; 269795; CAA93692.1; -.  
DR Hypothetical protein.  
KW SEQUENCE 132 AA; 15386 MW; 11233AF9B93D32F6 CRC64;


```

```

Query Match          100.0%; Score 19; DB 1; Length 132;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 AAPL 4
      1111
Db      67 AAPL 70

RESULT 26
CU3A_TENMO          STANDARD; PRT; 134 AA.
ID      CU3A_TENMO
AC      P80683;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Larval cuticle protein A3A (TM-A3A) (TM-LCP A3A).
OS      Tenebrio molitor (Yellow mealworm).
OC      Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC      Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC      Cucujiformia; Tenebrionidae; Tenebrio.
OX      NCBI_TaxID=7067;
RN      111
RP      SEQUENCE.
RC      TISSUE=Cuticle;
RX      MEDLINE=97218698; Pubmed=9066122;
RA      Andersen S.O., Rafn K., Roepstorff P.;
RT      "Sequence studies of proteins from larval and pupal cuticle of the
RL      Insect Biochem. Mol. Biol. 27:121-131(1997).
CC      -1- FUNCTION: COMPONENT OF THE CUTICLE OF THE LARVA OF TENEBRIO
CC      MOLITOR.
CC      -1- DOMAIN: THE TETRAPEPTIDE (A-A-P-[AV]) REPEATS FOUND THROUGHOUT THE
CC      PROTEIN ARE ALSO PRESENT IN MANY PROTEINS CONSTITUTING THE
CC      PROTECTIVE ENVELOPE OF OTHER SPECIES.
CC      -1- MASS SPECTROMETRY: MW=13974.3; METHOD=Electrospray.
CC      -1- SIMILARITY: CONTAINS 1 CUTICLE CONSENSUS DOMAIN.
DR      InterPro: IPR000618; Insect_cuticle.
DR      pfam: PF00379; Chitin_bind_4; 1.
DR      PRINTS: PR00947; CUTICLE.
DR      PROSITE: PS00233; CUTICLE; 1.
KW      Structural protein; Cuticle; Repeat.
FT      REPEAT 23 26 1.
FT      REPEAT 111 114 2.
SQ      SEQUENCE 134 AA; 13976 MW; F6AAAF06444FE93E CRC64;

Query Match          100.0%; Score 19; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 AAPL 4
      1111
Db      8 AAPL 11

RESULT 27
LAMP_PETMA          STANDARD; PRT; 139 AA.
ID      LAMP_PETMA
AC      P33575;
DT      01-FEB-1994 (Rel. 28, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Lamprin 0.9 precursor (Cartilage matrix protein).
OS      Petromyzon marinus (Sea lamprey).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Hyperoartia;
OC      Petromyzontiformes; Petromyzontidae; Petromyzon.
OX      NCBI_TaxID=7757;
RN      111
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 20-44.
RC      TISSUE=Cartilage;
RX      MEDLINE=93123269; Pubmed=7678258;
RA      Robson P., Wright G.M., Sitarz E., Malti A., Rawat M., Youson J.H.,

```

```

RA      Keeley F.W.;
RT      "Characterization of lamprin, an unusual matrix protein from lamprey
RT      cartilage. Implications for evolution, structure, and assembly of
RT      elastin and other fibrillar proteins.";
RL      J. Biol. Chem. 268:1440-1447(1993).
CC      -1- FUNCTION: SELF-AGGREGATING PROTEIN THAT MAKES PART OF THE SOLUBLE
CC      FORM OF LAMPRIN.
CC      -1- SUBUNIT: THE POLYMERIC LAMPRIN CHAINS SELF-AGGREGATE TO FORM
CC      FIBERS AND HAVE SECONDARY STRUCTURES PARTICULARLY RICH IN BETA-
CC      SHEETS AND IN BETA-TURNS.
CC      -1- SUBCELLULAR LOCATION: Extracellular matrix.
CC      -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 0.9-12 (SHOWN HERE) AND 0.9-10;
CC      SEEM TO BE PRODUCED BY ALTERNATIVE SPLICING.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use, by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; L05925; AAA49269.1; -.
DR      EMBL; L05924; AAA49268.1; -.
DR      PIR; C45051; C45051.
DR      PIR; A45051; A45051.
KW      Cartilage; Repeat; Connective tissue; Signal; Alternative splicing.
FT      SIGNAL 1 19
FT      CHAIN 20 139 LAMPRIN 0.9.
FT      DOMAIN 42 110 8 X 5 AA APPROXIMATE REPEATS.
FT      REPEAT 42 46 1.
FT      REPEAT 47 51 2.
FT      REPEAT 52 56 3.
FT      REPEAT 57 61 4.
FT      REPEAT 62 66 5.
FT      REPEAT 67 71 6.
FT      REPEAT 92 96 7.
FT      REPEAT 106 110 8.
FT      VARSPLIC 86 104 MISSING (IN ISOFORM 0.9-10).
SQ      SEQUENCE 139 AA; 13257 MW; E248ABB7A6CEC7C3 CRC64;

Query Match          100.0%; Score 19; DB 1; Length 139;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 AAPL 4
      1111
Db      129 AAPL 132

RESULT 28
EXBD_ECOLI          STANDARD; PRT; 141 AA.
ID      EXBD_ECOLI
AC      P18784;
DT      01-NOV-1990 (Rel. 16, Created)
DT      01-NOV-1990 (Rel. 16, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Biopolymer transport exbd protein.
GN      EXBD OR B3005 OR Z4358 OR ECS3889.
OS      Escherichia coli, and
OS      Escherichia coli O157:H7.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=562, 83334;
RN      111
RP      SEQUENCE FROM N.A.
RC      STRAIN-K12;
RX      MEDLINE=89359155; Pubmed=2670903;
RA      Eick-Helmerich K., Braun V.;
RT      "Import of biopolymers into Escherichia coli: nucleotide sequences of
RT      the exbB and exbD genes are homologous to those of the tolQ and tolR
RT      genes, respectively.";
RL      J. Bacteriol. 171:5117-5126(1989).

```

```

RN {2}
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=92278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
[5]
RN TOPOLOGY.
RP MEDLINE=92355532; PubMed=1644779;
RX Kampaenkel K., Braun V.;
RA "Membrane topology of the Escherichia coli Exbd protein.";
RT J. Bacteriol. 174:5485-5487(1992).
CC -i- FUNCTION: INVOLVED IN THE TONB-DEPENDENT ENERGY-DEPENDENT
CC TRANSPORT OF VARIOUS RECEPTOR-BOUND SUBSTRATES.
CC -i- SUBUNIT: THE ACCESSORY PROTEINS EXXB AND EXBD SEEM TO FORM A
CC COMPLEX WITH TONB.
CC -i- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
CC -i- SIMILARITY: BELONGS TO THE EXBD / TOLR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M28819; AAA23733.1; -
DR EMBL; U28377; AAA69172.1; -
DR EMBL; AE000383; AAC76041.1; -
DR EMBL; AE005530; AAG58141.1; -
DR EMBL; AP002563; BAB37312.1; -
DR PIR; JY0030; BVECED.
DR EcoGene; EG10272; exbd.
DR InterPro; IPR003400; Exbd.
DR Pfam; PF02472; Exbd; 1.
KW Transport; Protein transport; Transmembrane; Inner membrane;
KW Bacteriocin transport; Signal-anchor; Complete proteome.
FT DOMAIN 1 22 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 23 43 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT FT (PROBABLE).
FT FT PERIPLASMIC (PROBABLE).
SQ SEQUENCE 141 AA; 15527 MW; 25A539A2FAEB9F6C CRC64;
Query Match 100.0%; Score 19; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;

```

```

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPL 4
Db 37 AAPL 40
RESULT 29
ZRAP_ECOLI
ID ZRAP_ECOLI STANDARD; PRT; 141 AA.
AC P32682; Q8X6X2;
DT 01-OCT-1993 (Rel. 27, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Zinc resistance-associated protein precursor.
GN ZRAP OR ZRA OR B4002 OR Z5578 OR ECS4925.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes.";
RL Nucleic Acids Res. 21:5408-5417(1993).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
[4]
RN SEQUENCE OF 27-41, AND CHARACTERIZATION.
RP MEDLINE=98362023; PubMed=9694902;
RX Noll M., Petrukhin K., Lutsenko S.;
RT "Identification of a novel transcription regulator from Proteus
RT mirabilis, PMTR, revealed a possible role of YJAI protein in
RT balancing zinc in Escherichia coli.";
RL J. Biol. Chem. 273:21393-21401(1998).
CC -i- FUNCTION: Binds zinc. Could be an important component of the zinc-
CC balancing mechanism.
CC -i- SUBCELLULAR LOCATION: Periplasmic.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; U00006; AAC43100.1; ALT_INIT.
DR EMBL; AE000473; AAC76976.1; ALT_INIT.
DR EMBL; AE005631; AAC59199.1; ALT_INIT.
DR EMBL; AP002567; BAB38348.1; ALT_INIT.
DR EcoGene; EG11918; zrap.
KW Zinc; Periplasmic; Signal; Complete proteome.
FT SIGNAL 1 26
FT CHAIN 27 141
SQ SEQUENCE 141 AA; 15199 MW; 392A8B29CEBEC1B CRC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 1; Length 141;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
    ||||
Db 37 AAPL 40

RESULT 30
EXBD_PSEPU
ID EXBD_PSEPU STANDARD; PRT; 142 AA.
AC Q05606;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Biopolymer transport exbd protein.
GN EXBD.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WCS358;
RX MEDLINE=93172953; PubMed=8437515;
RA Bitter W., Tommassen J., Weisbeek P.J.;
RT "Identification and characterization of the exbB, exbD and tonB genes
RT of Pseudomonas putida WCS358: their involvement in
RT ferric-pseudobactin transport.";
RL Mol. Microbiol. 7:117-130(1993).
CC -|- FUNCTION: INVOLVED IN THE TONB-DEPENDENT ENERGY-DEPENDENT
CC TRANSPORT OF VARIOUS RECEPTOR-BOUND SUBSTRATES.
CC -|- SUBUNIT: THE ACCESSORY PROTEINS EXBB AND EXBD SEEM TO FORM A
CC COMPLEX WITH TONB.
CC -|- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane
CC (Probable).
CC -|- SIMILARITY: BELONGS TO THE EXBD / TOLR FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X70139; CAA49715.1; -.
DR PIR; S28443; S28443.
DR InterPro; IPR003400; Exbd.
DR Pfam; PF02472; Exbd; 1.
KW Transport; Protein transport; Transmembrane; Inner membrane.
FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 23 43 POTENTIAL.
FT DOMAIN 44 142 PERIPLASMIC (POTENTIAL).
SQ SEQUENCE 142 AA; 15225 MW; 1430086CBDCB935C CRC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 1; Length 142;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
    ||||

```

```

Db 37 AAPL 40

RESULT 31
SREL_PIG
ID SREL_PIG STANDARD; PRT; 144 AA.
AC 097676;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sterol regulatory element binding protein-1 (SREBP-1) (Sterol
DE regulatory element binding transcription factor 1) (Adipocyte
DE determination and differentiation-dependent factor 1) (ADD1)
DE (Fragment).
GN SREBP1 OR SREBP1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99410873; PubMed=10481259;
RA Ding S.T., McNeel R.L., Mersmann H.J.;
RT "Expression of porcine adipocyte transcripts: tissue distribution and
RT differentiation in vitro and in vivo.";
RL Comp. Biochem. Physiol. 123B:307-318(1999).
CC -|- FUNCTION: Transcriptional activator that binds to the sterol
CC regulatory element 1 (SRE-1) (5'-ATCACCACAC-3') Has dual sequence
CC specificity, binding to both an E-box motif (ATCAGCTGA) and to
CC SRE-1 (5'-ATCACCACAC-3'). Regulates the transcription of genes for
CC sterol biosynthesis and the LDL receptor gene (By similarity).
CC -|- SUBUNIT: Forms a tight complex with SCAP in the ER membrane.
CC Efficient DNA binding of the soluble transcription factor fragment
CC requires dimerization with another bHLH protein.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein that moves from
CC the endoplasmic reticulum to the Golgi in the absence of sterols
CC (By similarity).
CC -|- PTM: Under sterol-depleted conditions, SREBPs are cleaved
CC sequentially by site-1 and site-2 protease. The first cleavage
CC occurs within the luminal loop and is directly regulated by
CC sterols. The second cleavage by site-2 protease occurs within the
CC first transmembrane stretch and liberates the soluble
CC transcription factor. Cleavage by the cysteine proteases, caspase-
CC 3 and caspase-7, is induced during apoptosis, independent of
CC sterol levels (By similarity).
CC -|- MISCELLANEOUS: SREBPs have to be in a complex with the cleavage-
CC activating protein (SCAP) to move to the Golgi and be cleaved by
CC site-1 protease (By similarity).
CC -|- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AF102873; AAC78685.1; -.
DR HSSP; P36956; IAM9.
DR TRANSPAC; T04563; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH_1; PARTIAL.
DR PROSITE; PS50888; HLH_2; 1.
KW Transcription regulation; Activator; DNA-binding; Lipid metabolism;
KW Cholesterol metabolism; Nuclear protein; Transmembrane;
KW Endoplasmic reticulum; Golgi stack.
FT NON_TER 1 1
FT DOMAIN <1 144 CYTOPLASMIC (POTENTIAL).
FT DNA_BIND 96 109 BASIC DOMAIN.

```



```
FT DOMAIN 110 >144 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT NON_TER 144 144
SQ SEQUENCE 144 AA; 15222 MW; CBAFA771FA2745F1 CRC64;
Query Match 100.0%; Score 19; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
    1111
Db 51 AAPL 54

RESULT 32
SSI_STRAO STANDARD; PRT; 144 AA.
ID SSI_STRAO
AC P01006;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Subtilisin inhibitor precursor (SSI type).
GN SSI.
OS Streptomyces albogriseolus.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S-3253;
RX MEDLINE=89278057; Pubmed=2732212;
RA Obata S., Taguchi S., Kumagai I., Miura K.-I.;
RT "Molecular cloning and nucleotide sequence determination of gene
    encoding Streptomyces subtilisin inhibitor (SSI).";
RL J. Biochem. 105:367-371(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Miura K.-I., Kumagai I., Obata S., Kojima S., Taguchi S.;
RT "Partial alteration of a protein Streptomyces subtilisin inhibitor by
    site-directed mutagenesis.";
RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 64:147-149(1988).
RN [3]
RP SEQUENCE OF 32-144 FROM N.A.
RC STRAIN=S-3253;
RX MEDLINE=90128271; Pubmed=2482228;
RA Taguchi S., Nishiyama K.-I., Kumagai I., Miura K.-I.;
RT "Analysis of transcriptional control regions in the Streptomyces
    subtilisin-inhibitor-encoding gene.";
RL Gene 84:279-286(1989).
RN [4]
RP SEQUENCE OF 32-144 FROM N.A.
RX MEDLINE=93192316; Pubmed=8448204;
RA Ueda Y., Taguchi S., Nishiyama K.I., Kumagai I., Miura K.I.;
RT "Effect of a rare leucine codon, TTA, on expression of a foreign gene
    in Streptomyces lividans.";
RL Biochim. Biophys. Acta 1172:262-266(1993).
RN [5]
RP SEQUENCE OF 32-144.
RC STRAIN=S-3253;
RX MEDLINE=75133416; Pubmed=4376147;
RA Ikenaka T., Odani S., Sakai M., Nabeshima Y., Sato S., Mura S.;
RT "Amino acid sequence of an alkaline proteinase inhibitor
    (Streptomyces subtilisin inhibitor) from Streptomyces albogriseolus
    S-3253.";
RL J. Biochem. 76:1191-1209(1974).
RN [6]
RP REVISIONS TO 142-143.
RX MEDLINE=80227614; Pubmed=6993452;
RA Sakai M., Odani S., Ikenaka T.;
RT "Importance of the carboxyl-terminal four amino acid residues in the
    inhibitory activity of Streptomyces subtilisin inhibitor (with a
    revision of its carboxyl-terminal sequence).";
RL J. Biochem. 87:891-898(1980).
RN [7]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=79114456; Pubmed=763329;
RA Mitsui Y., Satow Y., Watanabe Y., Hirano S., Itaka Y.;
RT "Crystal structures of Streptomyces subtilisin inhibitor and its
    complex with subtilisin BPN'.";
RL Nature 277:447-452(1979).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=85033707; Pubmed=6387152;
RA Hirano S., Akagawa H., Mitsui Y., Itaka Y.;
RT "Crystal structure at 2.6-A resolution of the complex of subtilisin
    BPN' with streptomyces subtilisin inhibitor.";
RL J. Mol. Biol. 178:389-413(1984).
RN [9]
RP MUTAGENESIS OF MET-104.
RX MEDLINE=90365981; Pubmed=1366538;
RA Kojima S., Obata S., Kumagai I., Miura K.-I.;
RT "Alteration of the specificity of the Streptomyces subtilisin
    inhibitor by gene engineering.";
RL Biotechnology 8:449-452(1990).
RN [10]
RP MUTAGENESIS OF MET-104.
RX MEDLINE=91349173; Pubmed=1908859;
RA Kojima S., Nishiyama Y., Kumagai I., Miura K.-I.;
RT "Inhibition of subtilisin BPN' by reaction site P1 mutants of
    Streptomyces subtilisin inhibitor.";
RL J. Biochem. 109:377-382(1991).
CC -I- FUNCTION: STRONG INHIBITOR OF BACTERIAL SERINE PROTEASES SUCH AS
    SUBTILISIN.
CC -I- SUBUNIT: HOMODIMER.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- SIMILARITY: BELONGS TO THE SERINE PROTEASE INHIBITORS SSI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
    or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00402; BAA00305.1; -
DR EMBL; M31133; AAA26821.1; -
DR EMBL; M33134; AAA26822.1; -
DR EMBL; X73049; CAA51525.1; -
DR EMBL; M54887; AAA26827.1; -
DR PIR; JX0066; XSSMA.
DR PDB; 2SIC; 15-APR-93.
DR PDB; 3SIC; 31-JAN-94.
DR PDB; 5SIC; 31-JAN-94.
DR PDB; 3SSI; 17-AUG-96.
DR PDB; 2TLD; 15-JUL-92.
DR InterPro; IPR000691; Strep_subt_inhib.
DR Pfam; PF00720; SSI; 1.
DR PRINTS; PR00294; SSBTLNINHBR.
DR PRODOM; PD004028; Strep_subt_inhib; 1.
DR PROSITE; PS00999; SSI; 1.
KW Serine protease inhibitor; Signal; Repeat; 3D-structure.
FT SIGNAL 1 31
FT CHAIN 32 144 SUBTILISIN INHIBITOR.
FT REPEAT 33 37
FT REPEAT 39 43
FT DISULFID 66 81
FT DISULFID 102 132
FT ACY_SITE 104 105
FT SIMILAR 98 116
FT SIMILAR 129 132
FT MUTAGEN 104 104 WITH 12-29 OF BOVINE P.S.I.
FT MUTAGEN 104 104 WITH 32-35 OF BOVINE P.S.I.
FT MUTAGEN 104 104 M->K,R: ALSO INHIBITS TRYPSIN.
FT STRAND 104 104 M->Y,W: ALSO INHIBITS CHYMOTRYPSIN.
FT TURN 42 49 M->D,E,V,I,G,P: DECREASE IN INHIBITION.
FT TURN 53 54
FT STRAND 60 65
```



```

FT STRAND 71 73
FT TURN 75 76
FT HELIX 77 87
FT TURN 88 89
FT TURN 91 92
FT STRAND 100 103
FT STRAND 109 116
FT STRAND 121 128
FT HELIX 131 135
FT TURN 136 137
SQ SEQUENCE 144 AA; 14312 MW; BEA57AC7FDCD8004 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 19 AAPL 22

RESULT 33
SSI_STRCO STANDARD; PRT; 144 AA.
ID SSI_STRCO
AC P28591; Q9R644;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protease inhibitor precursor (SLPI) (Trypsin inhibitor STI1) (SILA-3).
GN STI1 OR SC00762 OR SCF81.21C.
OS Streptomyces coelicolor, and
OS Streptomyces lividans.
OC Bacteria; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902, 1916;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-85; 87-103 AND 105-143.
RC SPECIES=S.lividans;
RX MEDLINE=92147677; PubMed=1737780;
RA Strickler J.E., Berka T.R., Gorniak J., Fornwald J., Keys R.,
RA Rowland J.J., Rosenberg M., Taylor D.P.;
RT "Two novel Streptomyces protein protease inhibitors. Purification,
RT activity, cloning, and expression.";
RL J. Biol. Chem. 267:3236-3241(1992).
RN [2]
RP SEQUENCE OF 38-144.
RC SPECIES=S.lividans; STRAIN=66 / 1326;
RX MEDLINE=93015780; PubMed=1356971;
RA Ueda Y., Kojima S., Tsumoto K., Takeda S., Miura K.-I., Kumagai I.;
RT "A protease inhibitor produced by Streptomyces lividans 66 exhibits
RT inhibitory activities toward both subtilisin BPN' and trypsin.";
RL J. Biochem. 112:204-211(1992).
RN [3]
RP SEQUENCE OF 36-71.
RC SPECIES=S.lividans; STRAIN=66 / 1326;
RX MEDLINE=93222542; PubMed=7763545;
RA Taguchi S., Kikuchi H., Kojima S., Kumagai I., Nakase T., Miura K.-I.,
RA Momose H.;
RT "High frequency of SSI-like protease inhibitors among Streptomyces.";
RL Biosci. Biotechnol. Biochem. 57:522-524(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=S.coelicolor; STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabdinwitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
```

```

RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
RN [5]
RP SEQUENCE OF 38-144 FROM N.A.
RC SPECIES=S.coelicolor; STRAIN=A3(2);
RX MEDLINE=96162858; PubMed=8595854;
RA Taguchi S., Kojima S., Miura K., Momose H.;
RT "Taxonomic characterization of closely related Streptomyces spp. based
RT on the amino acid sequence analysis of protease inhibitor proteins.";
RL FEMS Microbiol. Lett. 135:169-173(1996).
CC -!- FUNCTION: STRONG INHIBITORY ACTIVITY TOWARD SUBTILISIN BPN' AND
CC TO A LESSER EXTENT, TO TRYPSIN.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SERINE PROTEASE INHIBITORS SSI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M80576; AAA26801.1; -
DR EMBL; AL133171; CAB61542.1; -
DR PIR; B42585; B42585.
DR PIR; JT0619; JT0619.
DR HSSP; P01006; 3SSI.
DR InterPro; IPR000691; Strep_subtl_inhib.
DR pfam; PF00720; SSI; 1.
DR PRINTS; PR00294; SSBTLINHBTR.
DR ProDom; PD004028; Strep_subtl_inhib; 1.
DR PROSITE; PS00999; SSI; 1.
KW Serine protease inhibitor; Signal; Complete proteome.
FT SIGNAL 1 35
FT CHAIN 36 144 PROTEASE INHIBITOR.
FT DISULFID 66 81
FT DISULFID 102 132
FT ACT_SITE 104 105 REACTIVE BOND.
SQ SEQUENCE 144 AA; 14433 MW; E748BADF078A4F8B CRC64;

Query Match 100.0%; Score 19; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 56 AAPL 59

RESULT 34
YE86_MYCPN STANDARD; PRT; 147 AA.
ID YE86_MYCPN
AC P75299;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN486 (P02_orf147).
GN MPN486 OR MP356.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Molllicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Hermann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000034; AAB96004.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 147 AA; 16257 MW; 0C15A7285A8B14C2 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 79 AAPL 82

RESULT 35
RL9_MYCPN STANDARD; PRT; 149 AA.
AC P75540;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L9.
GN RPLI OR MPN231 OR MP600.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
[1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- FUNCTION: BINDS TO THE 23S rRNA (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L9P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000058; AAB96248.1; -.
DR HSSP; P02417; 1DIV.
DR InterPro; IPR000244; Ribosomal_L9.
DR Pfam; PF01281; Ribosomal_L9; 1.
DR TIGRfams; TIGR00158; L9; 1.
DR PROSITE; PS00651; RIBOSOMAL_L9; FALSE_NEG.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 149 AA; 17152 MW; EED94AD27A4F08DF CRC64;

Query Match 100.0%; Score 19; DB 1; Length 149;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 36 AAPL 39

RESULT 36
EXON_PRIVN3 STANDARD; PRT; 150 AA.
ID EXON_PRIVN3
AC P30660;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Alkaline exonuclease (EC 3.1.11.-) (Fragment).
GN U112 OR AN.
OS Pseudorabies virus (strain NIA-3) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10349;
[1]
SEQUENCE FROM N.A.
RP MEDLINE=92365105; PubMed=1323689;
RX de Wind N., Domen J., Berns A.A.;
RA "Herpesviruses encode an unusual protein-serine/threonine kinase
RT which is nonessential for growth in cultured cells.";
RL J. Virol. 66:5200-5209(1992).
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES ALKALINE EXONUCLEASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M94870; AAA47482.1; -.
DR PIR; C42744; C42744.
DR InterPro; IPR001616; Herpes_alk_exo.
DR Pfam; PF01771; Herpes_alk_exo; 1.
KW Hydrolyase; Nuclease; Exonuclease.
FT NON_TER 150
FT SEQUENCE 150 AA; 15425 MW; 0433CB1A7CAED734 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 34 AAPL 37

RESULT 37
RL9_MYCPE STANDARD; PRT; 150 AA.
ID RL9_MYCPE
AC P47339;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L9.
GN RPLI OR RPL9 OR MG093.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
[1]
SEQUENCE FROM N.A.
RP STRAIN=ATCC 33530 / G-37;
RC MEDLINE=96026346; PubMed=7569993;
RX Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
CC -!- FUNCTION: BINDS TO THE 23S rRNA (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L9P FAMILY OF RIBOSOMAL PROTEINS.

```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39689; AAC71311.1; -.
DR HSSP; P02417; IDIV.
DR TIGR; MG093; -.
DR InterPro; IPR000244; Ribosomal_L9.
DR Pfam; PF01281; Ribosomal_L9; 1.
DR TIGRFAMs; TIGR00158; L9; 1.
DR PROSITE; PS00651; RIBOSOMAL_L9; FALSE_NEG.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 150 AA; 17380 MW; 50A19B6EAA48E8D6 CRC64;

Query Match      100.0%; Score 19; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 36 AAPL 39

RESULT 38
TTHY_CHICK
ID TTHY_CHICK STANDARD; PRT; 150 AA.
AC P27731;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transthyretin precursor (Prealbumin) (TBPA).
TTR.
GN Gallus gallus (Chicken).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-40.
RC STRAIN=White leghorn; TISSUE=Liver;
RX MEDLINE=92007844; PubMed=1833190;
RA Duan W., Achen M.G., Richardson S.J., Lawrence M.C.,
RA Wettenhall R.E.H., Jaworowski A., Schreiber G.;
RT "Isolation, characterization, cDNA cloning and gene expression of an
RT avian transthyretin. Implications for the evolution of structure and
RT function of transthyretin in vertebrates.";
RL Eur. J. Biochem. 200:679-687(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE=96195657; PubMed=8612621;
RA Sundé M., Richardson S.J., Chang L., Pettersson T.M., Schreiber G.,
RA Blake C.C.F.;
RT "The crystal structure of transthyretin from chicken.";
RL Eur. J. Biochem. 236:491-499(1996).
CC -1- FUNCTION: Thyroid hormone-binding protein. Probably transports
CC thyroxine from the bloodstream to the brain.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN THE CHOROID PLEXUS. ALSO
CC PRESENT IN THE LIVER.
CC -1- SIMILARITY: BELONGS TO THE TRANSTHYRETIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC -----
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60471; CAA43000.1; -.
DR PIR; S16035; S16035.
DR PIR; S17827; S17827.
DR PDB; 1TFP; 10-JUN-96.
DR InterPro; IPR000895; Transthyretin.
DR Pfam; PF00576; Transthyretin_1.
DR PRINTS; PR00189; TRNSTHYRETIN.
DR ProDom; PD003457; Transthyretin; 1.
DR SMART; SM00095; TR_THY; 1.
DR PROSITE; PS00768; TRANSTHYRETIN_1; 1.
DR PROSITE; PS00769; TRANSTHYRETIN_2; 1.
KW Albumin; Transport; Retinol-binding; Vitamin A; Thyroid hormone;
KW Signal; 3D-structure.
FT SIGNAL 1 20
FT CHAIN 21 150 TRANSTHYRETIN.
FT BINDING 38 38 THYROID HORMONES (BY SIMILARITY).
FT BINDING 77 77 THYROID HORMONES (BY SIMILARITY).
SQ SEQUENCE 150 AA; 16309 MW; 1F82A776A9996360 CRC64;

Query Match      100.0%; Score 19; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 20 AAPL 23

RESULT 39
TTHY_CROPO
ID TTHY_CROPO STANDARD; PRT; 150 AA.
AC O55245;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transthyretin precursor (Prealbumin).
TTR.
GN Crocodylus porosus (Crocodylle).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Crocodylinae; Crocodylus.
OX NCBI_TaxID=8502;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Prapunpoj P., Schreiber G.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Thyroid hormone-binding protein. Probably transports
CC thyroxine from the bloodstream to the brain (By similarity).
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE TRANSTHYRETIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ223148; CAA11129.1; -.
DR HSSP; P27731; 1TFP.
DR InterPro; IPR000895; Transthyretin.
DR Pfam; PF00576; Transthyretin; 1.
DR PRINTS; PR00189; TRNSTHYRETIN.
DR ProDom; PD003457; Transthyretin; 1.
DR SMART; SM00095; TR_THY; 1.
DR PROSITE; PS00768; TRANSTHYRETIN_1; 1.
DR PROSITE; PS00769; TRANSTHYRETIN_2; 1.
KW Signal; Albumin; Plasma; Transport; Thyroid hormone; Retinol-binding;
KW Vitamin A.
```

FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 150 TRANSTHYRETIN.
FT BINDING 38 38 THYROID HORMONES (BY SIMILARITY).
FT BINDING 77 77 THYROID HORMONES (BY SIMILARITY).
SQ SEQUENCE 150 AA; 16411 MW; F05DFAA6FBDE22A5 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
1111
Db 20 AAPL 23

Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AAPL 4
1111
Db 20 AAPL 23
Search completed: December 6, 2002, 13:28:34
Job time : 6.55556 secs

RESULT 40
TTHY_TILRU
ID TTHY_TILRU STANDARD; PRT; 150 AA.
AC P30623;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transthyretin precursor (Prealbumin) (TBPA).
GN TTR.
OS Tiliqua rugosa (Stumpy-tailed skink) (Stumpy-tailed lizard).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Scincomorpha; Scincoidae;
OC Scincidae; Tiliqua.
OX NCBI_TaxID=8527;
RN 11
RP SEQUENCE FROM N.A., AND 3D-STRUCTURE MODELING.
RC TISSUE-Brain;
RX MEDLINE=94057100; PubMed=8238627;
RA Achen M.G., Duan W., Peltersson T.M., Harms P.J., Richardson S.J.,
Lawrence M.C., Wettenhall R.E.H., Aldred A.R., Schreiber G.;
RT "Transthyretin gene expression in choroid plexus first evolved in
reptiles.";
RL Am. J. Physiol. 265:R982-R989(1993).
CC -1- FUNCTION: Thyroid hormone-binding protein. Probably transports
thyroxine from the bloodstream to the brain (By similarity).
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE CHOROID PLEXUS.
CC -1- SIMILARITY: BELONGS TO THE TRANSTHYRETIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X66697; CAA47238.1; -;
DR EMBL; M97509; AAA49620.1; -;
DR PIR; S25531; S25531.
DR HSP; P27731; ITP.
DR InterPro; IPR000895; Transthyretin.
DR Pfam; PF00576; Transthyretin; 1.
DR PRINTS; PR00189; TRANSTHYRETIN.
DR PRODOM; PD003457; Transthyretin; 1.
DR SMART; SM00095; TR_THY; 1.
DR PROSITE; PS00768; TRANSTHYRETIN_1; 1.
DR PROSITE; PS00769; TRANSTHYRETIN_2; 1.
KW Albumin; Transport; Retinol-binding; Vitamin A; Thyroid hormone;
KW Signal.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 150 TRANSTHYRETIN.
FT BINDING 38 38 THYROID HORMONES (BY SIMILARITY).
FT BINDING 77 77 THYROID HORMONES (BY SIMILARITY).
SQ SEQUENCE 150 AA; 16343 MW; 6C988E00949112EE CRC64;
Query Match 100.0%; Score 19; DB 1; Length 150;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:25:09 ; Search time 19.5556 Seconds
(without alignments)
42.146 Million cell updates/sec

Title: US-10-033-526-3
Perfect score: 19
Sequence: 1 AAPL 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 19 | 100.0 | 14 | 10 P82332 | P82332 pisum sativ |
| 2 | 19 | 100.0 | 16 | 4 Q9NN22 | Q9NN22 homo sapien |
| 3 | 19 | 100.0 | 18 | 4 Q9UCB2 | Q9UCB2 homo sapien |
| 4 | 19 | 100.0 | 21 | 4 Q14595 | Q14595 homo sapien |
| 5 | 19 | 100.0 | 25 | 2 Q9R4Q9 | Q9R4Q9 vibrio para |
| 6 | 19 | 100.0 | 32 | 16 Q9RTA1 | Q9RTA1 streptomyce |
| 7 | 19 | 100.0 | 38 | 16 Q8VKK3 | Q8VKK3 mycobacteri |
| 8 | 19 | 100.0 | 47 | 5 Q9VY69 | Q9VY69 drosophila |
| 9 | 19 | 100.0 | 48 | 16 Q8VKE9 | Q8VKE9 mycobacteri |
| 10 | 19 | 100.0 | 60 | 16 Q9PGC0 | Q9PGC0 xylella fas |
| 11 | 19 | 100.0 | 62 | 2 Q8RK84 | Q8RK84 mycobacteri |
| 12 | 19 | 100.0 | 62 | 16 Q9PCC8 | Q9PCC8 xylella fas |
| 13 | 19 | 100.0 | 63 | 4 Q96F47 | Q96F47 mus sapien |
| 14 | 19 | 100.0 | 63 | 11 Q925V0 | Q925V0 mus musculi |
| 15 | 19 | 100.0 | 63 | 16 Q9PNR9 | Q9PNR9 campylobact |
| 16 | 19 | 100.0 | 63 | 17 Q8ZV02 | Q8ZV02 pyrobaculum |

| | | | | | |
|----|----|-------|----|-----------|--------------------|
| 17 | 19 | 100.0 | 66 | 16 Q9ABK2 | Q9ABK2 caulobacter |
| 18 | 19 | 100.0 | 67 | 5 Q9BPG6 | Q9BPG6 conus texti |
| 19 | 19 | 100.0 | 67 | 10 Q9LIC5 | Q9LIC5 secale cere |
| 20 | 19 | 100.0 | 67 | 10 Q43018 | Q43018 prunus dulc |
| 21 | 19 | 100.0 | 67 | 12 Q81866 | Q81866 hepatitis e |
| 22 | 19 | 100.0 | 67 | 16 Q8ZBX4 | Q8ZBX4 yersinia pe |
| 23 | 19 | 100.0 | 68 | 6 Q29346 | Q29346 sus scrofa |
| 24 | 19 | 100.0 | 68 | 10 Q9FLN3 | Q9FLN3 arabidopsis |
| 25 | 19 | 100.0 | 68 | 16 Q8VJ02 | Q8VJ02 mycobacteri |
| 26 | 19 | 100.0 | 69 | 10 Q8S075 | Q8S075 oryza sativ |
| 27 | 19 | 100.0 | 69 | 12 Q9IPR8 | Q9IPR8 human enter |
| 28 | 19 | 100.0 | 69 | 12 Q9IPR7 | Q9IPR7 human enter |
| 29 | 19 | 100.0 | 69 | 12 Q9IPR6 | Q9IPR6 human enter |
| 30 | 19 | 100.0 | 69 | 12 Q9IPR5 | Q9IPR5 human enter |
| 31 | 19 | 100.0 | 69 | 12 Q9IPR4 | Q9IPR4 human enter |
| 32 | 19 | 100.0 | 69 | 12 Q9IPR3 | Q9IPR3 human enter |
| 33 | 19 | 100.0 | 69 | 12 Q9IPR2 | Q9IPR2 human enter |
| 34 | 19 | 100.0 | 69 | 12 Q9IPR1 | Q9IPR1 human enter |
| 35 | 19 | 100.0 | 69 | 12 Q9IPR0 | Q9IPR0 human enter |
| 36 | 19 | 100.0 | 69 | 12 Q9IPR0 | Q9IPR0 human enter |
| 37 | 19 | 100.0 | 69 | 12 Q9IPR0 | Q9IPR0 human enter |
| 38 | 19 | 100.0 | 69 | 12 Q9IPR0 | Q9IPR0 human enter |
| 39 | 19 | 100.0 | 69 | 12 Q9IPR0 | Q9IPR0 human enter |
| 40 | 19 | 100.0 | 69 | 12 Q9IPR0 | Q9IPR0 human enter |
| 41 | 19 | 100.0 | 69 | 12 Q9IPR0 | Q9IPR0 human enter |
| 42 | 19 | 100.0 | 69 | 12 Q9IPR0 | Q9IPR0 human enter |
| 43 | 19 | 100.0 | 69 | 12 Q9IPR0 | Q9IPR0 human enter |
| 44 | 19 | 100.0 | 69 | 12 Q9IPR0 | Q9IPR0 human enter |
| 45 | 19 | 100.0 | 69 | 12 Q9IPR0 | Q9IPR0 human enter |

ALIGNMENTS

RESULT 1
P82332 ID P82332 PRELIMINARY: PRT; 14 AA.
AC P82332;
DT 01-JUN-2000 (TREMBlrel. 14, Created)
DT 01-JUN-2000 (TREMBlrel. 14, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Unknown protein from 2D-page of thylakoid (SPOT115) (Fragments).
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_Taxid=3888;
RN [1]
RP SEQUENCE, SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
RC STRAIN=CV. DE GRACE; TISSUE=LEAF;
RX MEDLINE=20181728; PubMed=10715320;
RA Peltier J.-B., Friso G., Kalume D.E., Roepstorff P., Nilsson F.,
RA Adamska I., van Wijk K.J.;
RT "proteomics of the chloroplast: systematic identification and
RT targeting analysis of luminal and peripheral thylakoid proteins.";
RL Plant Cell 12:319-341(2000).
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE LUMEN OR
CC PERIPHERY.
CC -!- DEVELOPMENTAL STAGE: UNFOLDED AND FULLY DEVELOPED LEAVES.
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.5, ITS MW IS: 27.4 KDA.
KW Chloroplast; Thylakoid membrane.
FT NON_TER 1
FT NON_CONS 7
FT NON_TER 14
FT NON_CONS 14
SQ SEQUENCE 14 AA; 1461 MW; 0803373C9C937AAB CRC64;
Query Match 100.0%; Score 19; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPL 4
|||||

Db 10 AAPL 13

RESULT 2

Q9NNZ2

ID Q9NNZ2 PRELIMINARY; PRT; 16 AA.

AC Q9NNZ2; 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE Integrin alpha-2 subunit (Fragment).

GN ITGA2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98421383; PubMed=9746778;

RA Kritzik M., Savage B., Nugent D.J., Santoso S., Ruggeri Z.M.,

RA Kunicki T.J.;

RT "Nucleotide polymorphisms in the alpha2 gene define multiple alleles

RT that are associated with differences in platelet alpha2 beta1

RT density.";

RL Blood 92:2382-2388(1998).

DR EMBL; AF062039; AAF77577.1; -.

FT NON_TER

SQ SEQUENCE

16 AA; 1635 MW; F6A5FC64F83370A9 CRC64;

Query Match

Best Local Similarity 100.0%; Score 19; DB 4; Length 16;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4

Db 8 AAPL 11

RESULT 3

Q9UCB2

ID Q9UCB2 PRELIMINARY; PRT; 18 AA.

AC Q9UCB2; 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE Megakaryocyte potentiating factor (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE.

RX MEDLINE=94117486; PubMed=8288629;

RA Yamaguchi N., Hattori K., Oh-eda M., Kojima T., Imai N., Ochi N.;

RT "A novel cytokine exhibiting megakaryocyte potentiating activity from

RT a human pancreatic tumor cell line HPC-Y5.";

RL J. Biol. Chem. 269:805-808(1994).

SQ SEQUENCE

18 AA; 1726 MW; 8F309D27CAAF671 CRC64;

Query Match

Best Local Similarity 100.0%; Score 19; DB 4; Length 18;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4

Db 9 AAPL 12

RESULT 4

Q14595

ID Q14595 PRELIMINARY; PRT; 21 AA.

AC Q14595;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Integrin alpha 2 subunit (Fragment).

GN ITGA2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94103255; PubMed=8276836;

RA Zutter M.M., Santoro S.A., Painter A.S., Tsung Y.L., Gafford A.;

RT "The human alpha 2 integrin gene promoter. Identification of positive

RT and negative regulatory elements important for cell-type and

RT developmentally restricted gene expression.";

RL J. Biol. Chem. 269:463-469(1994).

DR EMBL; L24121; AAA16619.2; -.

FT NON_TER

SQ SEQUENCE

21 AA; 2134 MW; 2A9118BC2066A5FC CRC64;

Query Match

Best Local Similarity 100.0%; Score 19; DB 4; Length 21;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4

Db 8 AAPL 11

RESULT 5

Q9R4Q9

ID Q9R4Q9 PRELIMINARY; PRT; 25 AA.

AC Q9R4Q9; 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE CHITOVIBRIN (Fragment).

OS Vibrio parahaemolyticus.

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

OX NCBI_TaxID=670;

RN [1]

RP SEQUENCE.

RX MEDLINE=95210902; PubMed=7696854;

RA Gildemeister O.S., Zhu B.C., Laine R.A.;

RT "Chitovibrin: a chitin-binding lectin from Vibrio parahaemolyticus.";

RL Glycoconj. J. 11:518-526(1994).

SQ SEQUENCE

25 AA; 2719 MW; 0CBBADA5711A25CC CRC64;

Query Match

Best Local Similarity 100.0%; Score 19; DB 2; Length 25;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4

Db 4 AAPL 7

RESULT 6

Q9RIAI

ID Q9RIAI PRELIMINARY; PRT; 32 AA.

AC Q9RIAI; 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Hypothetical protein SCO0074.

GN SCO0074 OR SCJ11.03.

OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2) / M145;

```
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL109949; CAB52888.1; -.
KW Hypothetical protein.
SQ SEQUENCE 32 AA; 3399 MW; A3AAD1C42914D9E2 CRC64;

Query Match 100.0%; Score 19; DB 16; Length 32;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 5 AAPL 8

RESULT 7
O8VKK3 PRELIMINARY; PRT; 38 AA.
AC O8VKK3;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein MT0540.
GN MT0540.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE006954; AAK44763.1; -.
DR TIGR; MT0540; -.
KW Hypothetical protein.
SQ SEQUENCE 38 AA; 4269 MW; 2B21477D74C9820D CRC64;

Query Match 100.0%; Score 19; DB 16; Length 38;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 12 AAPL 15

RESULT 8
O9VY69 PRELIMINARY; PRT; 47 AA.
AC O9VY69;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CG15763 protein.
GN CG15763.
OS Drosophila melanogaster (Fruit fly).
```

```
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003494; AAF48336.1; -.
DR FlyBase; FBgn0040869; CG15763.
SQ SEQUENCE 47 AA; 5165 MW; CF33EE4280B70687 CRC64;

Query Match 100.0%; Score 19; DB 5; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 35 AAPL 38

RESULT 9
O8VKE9 PRELIMINARY; PRT; 48 AA.
AC O8VKE9;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein MT0771.
GN MT0771.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
```

RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / OSHKOSH;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE006968; AAK45007.1; -
 DR TIGR; MT0771; -
 KW Hypothetical protein.
 SQ SEQUENCE 48 AA; 4818 MW; 2AD26E6F349534CE CRC64;

Query Match 100.0%; Score 19; DB 16; Length 48;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
 Db 15 AAPL 18

RESULT 10
 O9PGC0 PRELIMINARY; PRT; 60 AA.
 AC O9PGC0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein Xf0382.
 GN Xf0382.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Lemos M.V., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montei-ro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhami A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira P.G., Rodrigues V., de Rosa A.J.M.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA de Souza A.P., Terezi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Mclanlis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";
 RL Nature 406:151-159(2000).
 DR EMBL; AE003890; AAF83192.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 60 AA; 6404 MW; BEDA0EA2BD737885 CRC64;

Query Match 100.0%; Score 19; DB 16; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
 Db 34 AAPL 37

RESULT 11
 O8RK84 PRELIMINARY; PRT; 62 AA.
 AC O8RK84;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative ferredoxin.
 GN MORB.
 OS Mycobacterium sp. RP1.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=156204;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RP1;
 RA Poupin P.;
 RT "Cloning and characterization of a gene encoding a cytochrome P450
 RT monooxygenase (MOR) involved in morpholine, piperidine, and
 RT pyrrolidine utilization in Mycobacterium sp. strain RP1: Isolation and
 RT partial characterization of the enzyme";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ310142; CAC84232.1; -
 SQ SEQUENCE 62 AA; 6793 MW; 13503FDD08A131B80 CRC64;

Query Match 100.0%; Score 19; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
 Db 19 AAPL 22

RESULT 12
 O9PCC8 PRELIMINARY; PRT; 62 AA.
 AC O9PCC8;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein Xf1853.
 GN Xf1853.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,

RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truiffi D., Tsai S.M., Tsuchiko M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vetore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.,
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"
RL Nature 406:151-159(2000).
DR EMBL; AE004006; AAF84659.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 62 AA; 7040 MW; 9456C4B05271D4AF CRC64;

Query Match 100.0%; Score 19; DB 16; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 6 AAPL 9

RESULT 13

ID Q96F47 PRELIMINARY; PRT; 63 AA.
AC Q96F47;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Similar to hypothetical protein MGC2668.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011620; AAH11620.1; -
SQ SEQUENCE 63 AA; 7143 MW; 96BA93664527546A CRC64;

Query Match 100.0%; Score 19; DB 4; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 2 AAPL 5

RESULT 14

ID Q925V0 PRELIMINARY; PRT; 63 AA.
AC Q925V0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE AMP activated protein kinase gamma 1 (Fragment).
GN PRKAG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-0 FROM N.A.
RC STRAIN=129/SV;

RX MEDLINE=21203559; PubMed=11306812;
RA Shamsadin R., Jantsan K., Adham I., Engel W.;
RT "Cloning, organisation, chromosomal localization and expression
RT analysis of the mouse *Prkag1* gene,"
RL Cytogenet. Cell Genet. 92:134-138(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Shamsadin R.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF266699; AAK49388.1;
DR EMBL; AF266698; AAK49388.1; JOINED.
DR MGI:108411; Prkag1.
KW Kinase.
FT NON_TER 1
FT NON_TER 63
SQ SEQUENCE 63 AA; 7103 MW; 8774FBB20C4C8C00 CRC64;

Query Match 100.0%; Score 19; DB 11; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 51 AAPL 54

RESULT 15

ID Q9PNR9 PRELIMINARY; PRT; 63 AA.
AC Q9PNR9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Putative periplasmic protein.
GN CJI021C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
RA Jags K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen *Campylobacter jejuni*
RT reveals hypervariable sequences,"
RL Nature 403:665-668(2000).
DR EMBL; AL139077; CAB73277.1; -
KW Complete proteome.
SQ SEQUENCE 63 AA; 7609 MW; 3408C6B9AD490D11 CRC64;

Query Match 100.0%; Score 19; DB 16; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 17 AAPL 20

RESULT 16

ID Q8ZV02 PRELIMINARY; PRT; 63 AA.
AC Q8ZV02;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein PAE2521a.

GN PAE2521A.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009881; AAL64254.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 63 AA; 6975 MW; DF4893406C8EC2AA CRC64;

Query Match 100.0%; Score 19; DB 17; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 10 AAPL 13

RESULT 17
Q9ABK2 PRELIMINARY; PRT; 66 AA.
AC Q9ABK2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein CC0225.
GN CC0225.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005697; AAK22212.1; -
DR TIGR; CC0225; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 66 AA; 7364 MW; 28EC010054E3C2FB CRC64;

Query Match 100.0%; Score 19; DB 16; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 11 AAPL 14

RESULT 18
Q9BPG6 PRELIMINARY; PRT; 67 AA.
AC Q9BPG6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Conotoxin scaffold IX.
OS Conus textile (Cloth-of-gold cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL; AF214962; AAG60390.1; -
SQ SEQUENCE 67 AA; 7597 MW; 7272AD7AD4D6BB3D CRC64;

Query Match 100.0%; Score 19; DB 5; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 61 AAPL 64

RESULT 19
Q9LLC5 PRELIMINARY; PRT; 67 AA.
AC Q9LLC5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE Ty3 retrotransposon structural protein (Fragment).
GN GAG.
OS Secale cereale (Rye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticaceae; Secale.
OX NCBI_TaxID=4550;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-TY3 CLASS RETROTRANSPOSON;
RX MEDLINE=20157063; PubMed=10655237;
RA Langdon T., Seago C., Jones R.N., Ougham H., Thomas H., Forster J.W.,
RA Jenkins G.;
RT "De novo evolution of satellite DNA on the rye B chromosome.";
RL Genetics 154:869-884(2000).
DR EMBL; AF223161; AAF79935.1; -
FT NON_TER 67
SQ SEQUENCE 67 AA; 7174 MW; AF070E6FBC9C7C8F CRC64;

Query Match 100.0%; Score 19; DB 10; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 14 AAPL 17

RESULT 20
Q43018 PRELIMINARY; PRT; 67 AA.
AC Q43018;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Lipid transfer protein.
OS Prunus dulcis (Almond) (Prunus amygdalus).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=3755;

RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=CV. TEXAS;
RA Suelves M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: X96715; CAA65476.1; -.
DR InterPro: IPR001768; Try/amyL_inhbltr.
DR Pfam: PF00234; tryp_alpha_amy1; 1.
SQ SEQUENCE 67 AA; 6909 MW; 88392E48CD4716B8 CRC64;

Query Match 100.0%; Score 19; DB 10; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 17 AAPL 20

RESULT 21

O81866 PRELIMINARY; PRT; 67 AA.
AC O81866;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Poly-proline hinge (Fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RA "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Rikihisa T., Mizuno K., Win K.M.;
RT "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92335008; PubMed=1630924;
RA Aye T.T., Uchida T., Ma X.Z., Iida F., Shikata T., Zhuang H.,
RA Win K.M.;
RT "Complete nucleotide sequence of a hepatitis E virus isolated from the
RT Xinjiang epidemic (1986-1988) of China.";
RL Nucleic Acids Res. 20:3512-3512(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92271462; PubMed=1589964;
RA Fry K.E., Tam A.W., Smith M.W., Kim J.P., Luk K.-C., Young L.M.,
RA Piatak M., Feldman R.A., Yun K.Y., Purdy M.A., McCaustland K.A.,
RA Bradley D.W., Reyes G.R.;
RT "Hepatitis E virus (HEV): Strain variation in the nonstructural gene
RT region encoding consensus motifs for an RNA-dependent RNA polymerase
RT and an ATP/GTP binding site.";
RL Virus Genes 6:173-185(1992).
RN [6]
RP SEQUENCE FROM N.A.

RX MEDLINE=93348763; PubMed=8346669;
RA Bi S.-L., Purdy M.A., McCaustland K.A., Margolis H.S., Bradley D.W.;
RT "The sequence of hepatitis E virus isolated directly from a single
RT source during an outbreak in China.";
RL Virus Res. 28:233-247(1993).
DR EMBL: L08816; AAA03185.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 67 AA; 6647 MW; BD7A2B36C6876EFD CRC64;

Query Match 100.0%; Score 19; DB 12; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 26 AAPL 29

RESULT 22

O8ZBX4 PRELIMINARY; PRT; 67 AA.
AC O8ZBX4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein YP03260.
GN YP03260.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL: AJ414156; CAC92494.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 67 AA; 7658 MW; 88FAB352A80CF57E CRC64;

Query Match 100.0%; Score 19; DB 16; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 47 AAPL 50

RESULT 23

O29346 PRELIMINARY; PRT; 68 AA.
AC O29346;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Fatty acid transporter protein (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;

RX MEDLINE=96327607; PubMed=8672129;
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine CDNA
library."
RL Mamm. Genome 7:509-517(1996).
DR EMBL; F14753; CAA23229.1; -
DR InterPro; IPR000566; Lipocln_cytrABP.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
FT NON_TER 1
FT KW 68
SQ SEQUENCE 68 AA; 7419 MW; ACCDFADB0178D8BA CRC64;

Query Match 100.0%; Score 19; DB 6; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 38 AAPL 41

RESULT 24

O9FLN3 PRELIMINARY; PRT; 68 AA.
AC O9FLN3;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE EmbiCAB62355.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98290546; PubMed=9628582;
RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
RT Sequence features of the regions of 1,456,315 bp covered by nineteen
RT physically assigned P1 and TAC clones."
RL DNA Res. 5:41-54(1998).
DR EMBL; AB010072; BAB09698.1; -
SQ SEQUENCE 68 AA; 7479 MW; 1E588234D88C3D8A CRC64;

Query Match 100.0%; Score 19; DB 10; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 5 AAPL 8

RESULT 25

O8VJ02 PRELIMINARY; PRT; 68 AA.
AC O8VJ02;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein MT3573.9.
GN MT3573.9.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE007161; AAK47923.1; -
DR TIGR; MT3573; -
KW Hypothetical protein.
SQ SEQUENCE 68 AA; 7604 MW; 8214FB5BA15DDDDF CRC64;

Query Match 100.0%; Score 19; DB 16; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 62 AAPL 65

RESULT 26

O8S075 PRELIMINARY; PRT; 69 AA.
AC O8S075;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE P0678F11.1 protein.
GN P0678F11.1.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0678F11."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003437; BAB86094.1; -
SQ SEQUENCE 69 AA; 7454 MW; 5A5BF82C44C65168 CRC64;

Query Match 100.0%; Score 19; DB 10; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPL 4
Db 3 AAPL 6

RESULT 27

O9IPS8 PRELIMINARY; PRT; 69 AA.
AC O9IPS8;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5033/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
RA Wang J., Shimada Y., Ishiko H.;
RT "Molecular epidemiology of enterovirus 71 in Taiwan.";

RL Arch. Virol. 146:589-600(2001).
DR EMBL; AB037250; BAB01494.1; -.
DR InterPro; IPR003138; Pico_P1A.
DR Pfam; PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match 100.0%; Score 19; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
Db 65 AAPL 68

RESULT 28
Q9IPS7 PRELIMINARY; PRT; 69 AA.
ID Q9IPS7
AC Q9IPS7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5142/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
Wang J., Shimada Y., Ishiko H.;
RT "Molecular epidemiology of enterovirus 71 in Taiwan."
RL Arch. Virol. 146:589-600(2001).
DR EMBL; AB037251; BAB01495.1; -.
DR InterPro; IPR003138; Pico_P1A.
DR Pfam; PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match 100.0%; Score 19; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
Db 65 AAPL 68

RESULT 29
Q9IPS6 PRELIMINARY; PRT; 69 AA.
ID Q9IPS6
AC Q9IPS6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H0106/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
Wang J., Shimada Y., Ishiko H.;
RT "Molecular epidemiology of enterovirus 71 in Taiwan."
RL Arch. Virol. 146:589-600(2001).
DR EMBL; AB037252; BAB01496.1; -.
DR InterPro; IPR003138; Pico_P1A.

DR Pfam; PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match 100.0%; Score 19; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
Db 65 AAPL 68

RESULT 30
Q9IPS5 PRELIMINARY; PRT; 69 AA.
ID Q9IPS5
AC Q9IPS5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1657/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
Wang J., Shimada Y., Ishiko H.;
RT "Molecular epidemiology of enterovirus 71 in Taiwan."
RL Arch. Virol. 146:589-600(2001).
DR EMBL; AB037253; BAB01497.1; -.
DR InterPro; IPR003138; Pico_P1A.
DR Pfam; PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match 100.0%; Score 19; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
||||
Db 65 AAPL 68

RESULT 31
Q9IPS4 PRELIMINARY; PRT; 69 AA.
ID Q9IPS4
AC Q9IPS4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1658/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
Wang J., Shimada Y., Ishiko H.;
RT "Molecular epidemiology of enterovirus 71 in Taiwan."
RL Arch. Virol. 146:589-600(2001).
DR EMBL; AB037254; BAB01498.1; -.
DR InterPro; IPR003138; Pico_P1A.
DR Pfam; PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match 100.0%; Score 19; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPL 4
Db 65 AAPL 68

RESULT 32
Q9IPS3
ID Q9IPS3 PRELIMINARY; PRT; 69 AA.
AC Q9IPS3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Polypeptide (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1569/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
RA Wang J., Shmada Y., Ishiko H.;
RT "Molecular epidemiology of enterovirus 71 in Taiwan.";
RL Arch. Virol. 146:589-600(2001).
DR EMBL; AB037255; BAB01499.1; -.
DR InterPro; IPR003138; Pico_P1A.
DR Pfam; PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match 100.0%; Score 19; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPL 4
Db 65 AAPL 68

RESULT 33
Q9IPS2
ID Q9IPS2 PRELIMINARY; PRT; 69 AA.
AC Q9IPS2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Polypeptide (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1499/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
RA Wang J., Shmada Y., Ishiko H.;
RT "Molecular epidemiology of enterovirus 71 in Taiwan.";
RL Arch. Virol. 146:589-600(2001).
DR EMBL; AB037256; BAB01500.1; -.
DR InterPro; IPR003138; Pico_P1A.
DR Pfam; PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match 100.0%; Score 19; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPL 4
Db 65 AAPL 68

RESULT 34
Q9IPS1
ID Q9IPS1 PRELIMINARY; PRT; 69 AA.
AC Q9IPS1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Polypeptide (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=480/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
RA Wang J., Shmada Y., Ishiko H.;
RT "Molecular epidemiology of enterovirus 71 in Taiwan.";
RL Arch. Virol. 146:589-600(2001).
DR EMBL; AB037257; BAB01501.1; -.
DR InterPro; IPR003138; Pico_P1A.
DR Pfam; PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match 100.0%; Score 19; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPL 4
Db 65 AAPL 68

RESULT 35
Q9IPS0
ID Q9IPS0 PRELIMINARY; PRT; 69 AA.
AC Q9IPS0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Polypeptide (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=874/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
RA Wang J., Shmada Y., Ishiko H.;
RT "Molecular epidemiology of enterovirus 71 in Taiwan.";
RL Arch. Virol. 146:589-600(2001).
DR EMBL; AB037258; BAB01502.1; -.
DR InterPro; IPR003138; Pico_P1A.
DR Pfam; PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match 100.0%; Score 19; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4

Db 65 AAPL 68

RESULT 36

Q91PR9 PRELIMINARY; PRT; 69 AA.
AC Q91PR9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Polypeptide (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=737/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
RT "Molecular epidemiology of enterovirus 71 in Taiwan.";
RL Arch. Virol. 146:589-600(2001).
DR EMBL: AB037259; BAB01503.1; -;
DR InterPro: IPR003138; Pico_P1A.
DR Pfam: PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match

Best Local Similarity 100.0%; Score 19; DB 12; Length 69;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 65 AAPL 68

RESULT 37

Q91PR8 PRELIMINARY; PRT; 69 AA.
AC Q91PR8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Polypeptide (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=607/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
RT "Molecular epidemiology of enterovirus 71 in Taiwan.";
RL Arch. Virol. 146:589-600(2001).
DR EMBL: AB037260; BAB01504.1; -;
DR InterPro: IPR003138; Pico_P1A.
DR Pfam: PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match

Best Local Similarity 100.0%; Score 19; DB 12; Length 69;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 65 AAPL 68

RESULT 38

Q91PR7 PRELIMINARY; PRT; 69 AA.
AC Q91PR7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Polypeptide (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=602/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
RT "Molecular epidemiology of enterovirus 71 in Taiwan.";
RL Arch. Virol. 146:589-600(2001).
DR EMBL: AB037261; BAB01505.1; -;
DR InterPro: IPR003138; Pico_P1A.
DR Pfam: PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match

Best Local Similarity 100.0%; Score 19; DB 12; Length 69;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 65 AAPL 68

RESULT 39

Q91PR6 PRELIMINARY; PRT; 69 AA.
AC Q91PR6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Polypeptide (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=588/98;
RX MEDLINE=21237062; PubMed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
RT "Molecular epidemiology of enterovirus 71 in Taiwan.";
RL Arch. Virol. 146:589-600(2001).
DR EMBL: AB037262; BAB01506.1; -;
DR InterPro: IPR003138; Pico_P1A.
DR Pfam: PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match

Best Local Similarity 100.0%; Score 19; DB 12; Length 69;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 65 AAPL 68

RESULT 40

Q91PR5

ID Q9IPR5 PRELIMINARY; PRT; 69 AA.
AC Q9IPR5;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Human enterovirus 71.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_Taxid=39054;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1226/98;
RX MEDLINE=21237062; Pubmed=11338392;
RA Chu P.-Y., Lin K.-H., Hwang K.-P., Chou L.-C., Wang C.-F., Shih S.-R.,
RA Wang J., Shmada Y., Ishiko H.;
RT "Molecular epidemiology of enterovirus 71 in Taiwan."
RL Arch. Virol. 146:589-600(2001).
DR EMBL: AB037263; BAB01507.1; -
DR InterPro; IPR003138; Pico_P1A.
DR Pfam; PF02226; Pico_P1A; 1.
FT NON_TER 69
SQ SEQUENCE 69 AA; 7493 MW; 69B04974EE01BD4B CRC64;

Query Match 100.0%; Score 19; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPL 4
Db 65 AAPL 68

Search completed: December 6, 2002, 13:30:09
Job time : 20.5556 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:19:54 ; Search time 39 Seconds
(without alignments)
20.500 Million cell updates/sec

Title: US-10-033-526-4
Perfect score: 29
Sequence: 1 AAAAPF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Match | Query length | DB ID | Description |
|------------|-------|---------|--------------|-------|-------------|
| 1 | 29 | 100.0 | 6 | 23 | AAO18046 |
| 2 | 29 | 100.0 | 18 | 23 | AAU70970 |
| 3 | 29 | 100.0 | 53 | 20 | AAV11554 |
| 4 | 29 | 100.0 | 55 | 17 | AAU98596 |
| 5 | 29 | 100.0 | 102 | 23 | AAU70953 |
| 6 | 29 | 100.0 | 140 | 22 | AAU47148 |
| 7 | 29 | 100.0 | 362 | 19 | AAW59495 |
| 8 | 29 | 100.0 | 362 | 20 | AAV08994 |
| 9 | 29 | 100.0 | 362 | 22 | AAW52288 |
| 10 | 29 | 100.0 | 389 | 19 | AAW52823 |

| | | | | | | |
|----|----|-------|-----|----|----------|--------------------|
| 11 | 29 | 100.0 | 397 | 22 | AAU90077 | C glutamicum prote |
| 12 | 29 | 100.0 | 397 | 22 | AAU79052 | Corynebacterium gl |
| 13 | 29 | 100.0 | 405 | 23 | AAU51979 | Barley biotin synt |
| 14 | 29 | 100.0 | 410 | 22 | AAU70694 | Mouse DEC2a protei |
| 15 | 29 | 100.0 | 482 | 22 | AAU70692 | Human DEC2a protei |
| 16 | 29 | 100.0 | 484 | 22 | AAU70693 | Human DEC2b protei |
| 17 | 29 | 100.0 | 618 | 19 | AAW52827 | Human chromosome X |
| 18 | 29 | 100.0 | 723 | 22 | AAU62129 | Drosophila melanog |
| 19 | 29 | 100.0 | 747 | 22 | AAU69083 | Drosophila melanog |
| 20 | 29 | 100.0 | 860 | 21 | AAU23271 | Balanus amphitrite |
| 21 | 26 | 89.7 | 9 | 18 | AAU43821 | Specific human leu |
| 22 | 26 | 89.7 | 9 | 18 | AAU43809 | Specific human leu |
| 23 | 26 | 89.7 | 14 | 23 | AAU99695 | Peptide encoded by |
| 24 | 26 | 89.7 | 83 | 22 | AAU53233 | Propionibacterium |
| 25 | 26 | 89.7 | 138 | 21 | AAU32830 | Eucalyptus grandis |
| 26 | 26 | 89.7 | 159 | 22 | AAU09853 | Novel human diagno |
| 27 | 26 | 89.7 | 204 | 18 | AAU33276 | S. fradiae tylosin |
| 28 | 26 | 89.7 | 210 | 13 | AAU22377 | Antigen mc-4c. E1 |
| 29 | 26 | 89.7 | 325 | 22 | AAU31087 | C glutamicum prote |
| 30 | 26 | 89.7 | 325 | 22 | AAU79630 | Corynebacterium gl |
| 31 | 26 | 89.7 | 327 | 21 | AAU31744 | Arabidopsis thalia |
| 32 | 26 | 89.7 | 329 | 21 | AAU30330 | Arabidopsis thalia |
| 33 | 26 | 89.7 | 331 | 22 | AAU12782 | Novel human diagno |
| 34 | 26 | 89.7 | 407 | 23 | AAU10957 | Even-skipped homeo |
| 35 | 26 | 89.7 | 407 | 23 | AAU10958 | Even-skipped homeo |
| 36 | 26 | 89.7 | 407 | 23 | AAU10959 | Even-skipped homeo |
| 37 | 26 | 89.7 | 414 | 22 | AAU78989 | Escherichia coli L |
| 38 | 26 | 89.7 | 416 | 23 | AAU92120 | Herbicidally activ |
| 39 | 26 | 89.7 | 431 | 21 | AAU19410 | An aromatic class |
| 40 | 26 | 89.7 | 431 | 23 | AAU81703 | Arabidopsis sp. AT |
| 41 | 26 | 89.7 | 431 | 23 | AAU72773 | Arabidopsis thali |
| 42 | 26 | 89.7 | 434 | 21 | AAU31743 | Arabidopsis thalia |
| 43 | 26 | 89.7 | 440 | 22 | AAU69006 | Drosophila melanog |
| 44 | 26 | 89.7 | 460 | 23 | AAU92909 | Herbicidally activ |
| 45 | 26 | 89.7 | 463 | 22 | AAU59725 | Drosophila melanog |

ALIGNMENTS

| | | |
|----------|--|-----------------------------------|
| RESULT 1 | AAO18046 | AAO18046 standard; peptide; 6 AA. |
| ID | AAO18046; | |
| AC | AAO18046; | |
| XX | | |
| DT | 30-AUG-2002 (first entry) | |
| XX | | |
| DE | C-terminal truncated apoE formation inhibitor peptide #4. | |
| XX | | |
| KW | C-terminal truncated apoE; apolipoprotein E; apoE; Alzheimer's disease; | |
| KW | neurofibrillary tangle; apoE4 allele; neurotrophic; neuroprotective; | |
| KW | cardiant; vulnary; cerebroprotective; coronary artery disease; | |
| KW | head trauma; stroke. | |
| XX | | |
| OS | Synthetic. | |
| XX | | |
| PN | WO200238108-A2. | |
| XX | | |
| PD | 16-MAY-2002. | |
| XX | | |
| PF | 02-NOV-2001; 2001WO-US51172. | |
| XX | | |
| PR | 03-NOV-2000; 2000US-245737P. | |
| XX | | |
| PA | (GLAD-) GLADSTONE INST J DAVID. | |
| XX | | |
| PI | Huang Y, Mahley RW; | |
| XX | | |
| DR | WPT; 2002-490051/52. | |
| XX | | |
| PT | Inhibiting neurofibrillary tangles formation, useful for treating e.g. Alzheimer's, coronary artery disease or stroke, by reducing the | |

PT formation of carboxyl-terminal truncated form of apolipoprotein E in a
PT neuron of the individual -
XX
PS Claim 24; Page 62; 75pp; English.
XX
CC The present invention relates to a method of inhibiting the formation of
CC neurofibrillary tangles in an individual, which involves reducing the
CC formation of a carboxyl-terminal truncated form of apolipoprotein E
CC (apoE) in a neurone in the individual. The method is useful for
CC inhibiting the formation of neurofibrillary tangles in an individual. The
CC reduction in the formation of carboxyl-terminal truncated apoE treats a
CC disorder related to apoE in an individual, specifically Alzheimer's
CC disease, coronary artery disease, head trauma or stroke. The present
CC sequence is a peptide capable of inhibiting the formation of the
CC C-terminal truncated form of apoE.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 1 AAAAPF 6

RESULT 2
AAU70970
ID AAU70970 standard; Peptide; 18 AA.
XX
AC AAU70970;
XX
DT 25-FEB-2002 (first entry)
XX
DE M. tuberculosis Rv1386 protein immunogenic peptide P3.
XX
KW Tuberculosis; Tuberculostatic; antibacterial; vaccine; Rv0284;
KW Rv0285; Rv0455c; Rv0569; Rv1195; Rv1386; Rv3477; Rv3878; Rv3879;
KW MT3106.1; ORF13A; Rv0284ct; Mycobacterium bovis; Mycobacterium africanum;
KW BCG vaccine; immunogenic peptide.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200179274-A2.
XX
PD 25-OCT-2001.
XX
PF 19-APR-2001; 2001WO-DK00276.
XX
PR 19-APR-2000; 2000DK-0000666.
PR 21-FEB-2001; 2001DK-0000283.
XX
PA (STAT-) STATENS SERUM INST.
XX
PI Agger EM, Andersen P, Okkels LMM, Weldingh K;
XX
DR WPI; 2002-061970/08.
XX
PT New Mycobacterium tuberculosis antigens, useful for diagnosing
PT tuberculosis, and as a vaccine for treating or preventing infections
PT caused by species of tuberculosis complex -
XX
PS Example 3; Page 40; 111pp; English.
XX
CC The invention relates to a substantially pure polypeptide comprising
CC an amino acid sequence selected from Rv0284, Rv0285, Rv0455c,
CC Rv0569, Rv1195, Rv1386, Rv3477, Rv3878, Rv3879c or MT3106.1 (also
CC disclosed are ORF13A and Rv0284ct), or their immunogenic portion, nucleic
CC acids encoding them and an amino acid sequence analogue having at least
CC 70% sequence identity to the polypeptide and is immunogenic.
CC The protein is useful in preparing a pharmaceutical composition for
CC diagnosing tuberculosis and in preparing a vaccine against tuberculosis

CC caused by virulent mycobacteria. The vaccine or immunogenic/
CC pharmaceutical composition can be used prophylactically in a subject not
CC infected with a virulent mycobacterium, or therapeutically in a subject
CC already infected with a virulent mycobacterium. The protein is useful for
CC preventing, treating and detecting infections caused by species of
CC tuberculosis complex (M. tuberculosis, M. bovis, M. africanum). The
CC nucleic acids may be used for effecting in vivo expression of the
CC antigen, and in diagnostic assays for detecting the presence of
CC pathogenic organisms in a sample. The vaccine is an improvement of the
CC living BCG vaccine presently available, where one or more copies of the
CC DNA sequence encoding one or more polypeptide has been incorporated into
CC the genome of the microorganism to allow the microorganism to express and
CC secrete the polypeptide. Incorporation of more than one copy of a
CC nucleotide sequence enhances the immune response. The present
CC sequence represents an immunogenic peptide derived from an M.
CC tuberculosis protein of the invention.
XX
SQ Sequence 18 AA;

Query Match 100.0%; Score 29; DB 23; Length 18;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 9 AAAAPF 14

RESULT 3
AAV11554
ID AAV11554 standard; Protein; 53 AA.
XX
AC AAV11554;
XX
DT 16-JUN-1999 (first entry)
XX
DE Human 5' EST secreted protein SEQ ID NO:206.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
OS Homo sapiens.
XX
PN WO9906439-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB01233.
XX
PR 01-AUG-1997; 97US-0904468.
XX
PA (GEST) GENSET.
XX
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
DR WPI; 1999-153700/13.
DR N-PSDB; AAX40272.
XX
PT New nucleic acids encoding human secreted proteins - obtained from
PT CDNA libraries derived from liver, lung, large intestine, colon,
PT thyroid and pancreas tissue
XX
PS Claim 27; Page 320-321; 398pp; English.
XX
CC AAX40251 to AAX40397 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAV11533 to
CC AAV11679, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They

CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 53 AA;

Query Match 100.0%; Score 29; DB 20; Length 53;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
111111
Db 17 AAAAPF 22

RESULT 4
AAR98596
ID AAR98596 standard; peptide; 55 AA.

XX AAR98596;

DT 07-FEB-1997 (first entry)

DE Human CD14 mutant 3C-10 MAb epitope (SCD14 (45-48)A).

XX Antlinflammatory; lipopolysaccharide; LPS; Gram-positive bacteria;

KW CD14; treatment; septic shock; inflammatory bowel disease;

KW liver failure; graft versus host disease; pancreatitis; tuberculosis;

KW adult respiratory distress syndrome; detection; quantification.

XX Homo sapiens.

OS

XX WO9620956-A1.

PD 11-JUL-1996.

XX 28-DEC-1995; 95WO-US16977.

PR 30-DEC-1994; 94US-0366953.

XX (AMGE-) AMGEN INC.

PA (UYRO) UNIV ROCKEFELLER.

PI Juan S, Lichenstein HS, Narhi LO, Wright SD;

DR WPI; 1996-333943/33.

XX Peptide(s) based on CD14 sequences - which bind to
PT lipo:poly:saccharide and inhibit CD14 mediated inflammatory
PT responses

XX Example 9; Figure 11; 103pp; English.

XX Antinflammatory peptides based on the human CD14 sequence (See
CC AAR98570-75 and AAR98577-80) can bind to lipopolysaccharide (LPS) and
CC inhibit binding of LPS or Gram positive cell components to CD14,
CC thus reducing or eliminating CD14 mediated inflammatory responses.
CC They can be used for treating inflammatory conditions in particular,
CC septic shock, inflammatory bowel disease, acute and chronic liver
CC failure, graft versus host disease, intestinal or liver transplant,
CC adult respiratory distress syndrome, acute pancreatitis and
CC tuberculosis. They can also be used to remove, detect or quantitate
CC LPS or Gram-positive cell components in samples. The peptides are
CC used in doses of 0.1-100 mg/kg by parenteral or oral routes. A

CC series of CD14 mutants were used to localise and characterise an
CC epitope from CD14 specific for the monoclonal antibody 3C-10 (The
CC mutant peptide epitopes are described in AAR98590-97). To characterise
CC the 3C-10 epitope, site directed mutagenesis was used to create a
CC series of cDNAs encoding soluble CD14 having alanine substituted at
CC various positions between amino acids 1-55. The ability of 3C-10 to
CC bind to these mutant proteins was tested using a BIAcore biosensor
CC instrument. The epitope was localised to amino acids 7-14 inclusive
CC of CD14.

XX Sequence 55 AA;

Query Match 100.0%; Score 29; DB 17; Length 55;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
111111
Db 45 AAAAPF 50

RESULT 5
AAU70953
ID AAU70953 standard; Protein; 102 AA.

XX AAU70953;

DT 25-FEB-2002 (first entry)

DE M. tuberculosis Rv1386 protein.

XX Tuberculosis; Tuberculostatic; antibacterial; vaccine; Rv0284;

KW Rv0285; Rv0455c; Rv0569; Rv1195; Rv1386; Rv3477; Rv3878; Rv3879;

KW MT3106.1; ORF13A; Rv0284ct; Mycobacterium bovis; Mycobacterium africanum;

KW BCG vaccine.

XX Mycobacterium tuberculosis.

OS WO200179274-A2.

XX 25-OCT-2001.

PD 19-APR-2001; 2001WO-DK00276.

XX 19-APR-2000; 2000DK-0000666.

PR 21-FEB-2001; 2001DK-0000283.

XX (STAT-) STATENS SERUM INST.

XX Agger EM, Andersen P, Okkels LMM, Weldingh K;

PI WPI; 2002-061970/08.

DR N-PSDB; AAS95790.

XX New Mycobacterium tuberculosis antigens, useful for diagnosing
PT tuberculosis, and as a vaccine for treating or preventing infections
PT caused by species of tuberculosis complex

XX Claim 1; Page 77; 11pp; English.

XX The invention relates to a substantially pure polypeptide comprising
CC an amino acid sequence selected from Rv0284, Rv0285, Rv0455c,
CC Rv0569, Rv1195, Rv1386, Rv3477, Rv3878, Rv3879c or MT3106.1 (also
CC disclosed are ORF13A and Rv0284ct), or their immunogenic portion, nucleic
CC acids encoding them and an amino acid sequence analogue having at least
CC 70% sequence identity to the polypeptide and is immunogenic.
CC The protein is useful in preparing a pharmaceutical composition for
CC diagnosing tuberculosis and in preparing a vaccine against tuberculosis
CC caused by virulent mycobacteria. The vaccine or immunogenic/
CC pharmaceutical composition can be used prophylactically in a subject not
CC infected with a virulent mycobacterium, or therapeutically in a subject
CC already infected with a virulent mycobacterium. The protein is useful for
CC preventing, treating and detecting infections caused by species of

CC tuberculosis complex (M. tuberculosis, M. bovis, M. africanum). The
CC nucleic acids may be used for effecting in vivo expression of the
CC antigen, and in diagnostic assays for detecting the presence of
CC pathogenic organisms in a sample. The vaccine is an improvement of the
CC living BCG vaccine presently available, where one or more copies of the
CC DNA sequence encoding one or more polypeptide has been incorporated into
CC the genome of the microorganism to allow the microorganism to express and
CC secrete the polypeptide. Incorporation of more than one copy of a
CC nucleotide sequence enhances the immune response. The present
CC sequence represents an M. tuberculosis protein of the invention.
SQ Sequence 102 AA;

Query Match 100.0%; Score 29; DB 23; Length 102;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
 |||||
Db 29 AAAAPF 34

RESULT 6
AAU47148
ID AAU47148 standard; Protein; 140 AA.
XX
AC AAU47148;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #8044.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59537.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 8343; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention

CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 140 AA;

Query Match 100.0%; Score 29; DB 22; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
 |||||
Db 20 AAAAPF 25

RESULT 7
AAW59495
ID AAW59495 standard; Protein; 302 AA.
XX
AC AAW59495;
XX
DT 08-SEP-1998 (first entry)
XX
DE Human TFE3 5'-RACE fusion product from cell line UOK109.
XX
KW PRCG; papillary renal cell carcinoma; TFE3; transcription factor; RACE;
KW fusion protein; translocation; diagnosis; treatment; Nono; p54-nrb.
XX
OS Homo sapiens.
OS Synthetic.
OS
FH Key Location/Qualifiers
FH Protein 1..302
FT /label= TFE3
FT /note= "partial transcription factor sequence"
XX
PN WO9806871-A1.
XX
PD 19-FEB-1998.
XX
PF 13-AUG-1997; 97WO-GB02209.
XX
PR 13-AUG-1996; 96GB-0016986.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX
PI Clark J, Cooper C, Shipley J;
XX
DR WPI; 1998-159557/14.
DR N-PSDB; AAV20961.
XX
PT Diagnosing papillary renal cell carcinoma by detecting gene
PT trans-location - resulting in fusion of TFE3 gene with some other
PT gene, also related vectors, transformed cells, specific binding
PT reagents, peptide(s) encoded by fusions and therapeutic anti-sense
PT sequences
XX
PS Claim 9; Fig 4B; 71pp; English.
XX
CC This sequence represents a partial transcription factor, TFE3 sequence
CC from a 5'-RACE product derived from a novel fusion protein
CC constructed from the Nono protein (also known as p54-nrb) and the
CC transcribed factor TFE3 which is found in cell line UOK109 and used in
CC a method for the diagnosis, prophylactic and therapeutic treatment of
CC papillary renal cell carcinoma. The translocation t(X;1) (p11.2;q21.2)
CC found in papillary renal cell carcinoma (PRCC) associated protein (PRCC)

CC results in a fusion of the TFE3 gene with a new chromosome 1 gene
CC designated PRCC (at 1q21.2), resulting in expression of a fusion protein
CC between the N-terminus of PRCC and almost the whole of the TFE3 gene.
CC Normal TFE3 transcripts are no longer produced. Two other fusion partners
CC for TFE3 have also been detected: Nono, from a invX (p11.2; q13-24 or 12)
CC translocation and the PSF splice factor gene, resulting in t(X;1)
CC (p11.2;p34). These trans-locations define a subgroup of PRCC generally
CC encountered in patients younger than 25.

XX
SQ Sequence 302 AA;

Query Match 100.0%; Score 29; DB 19; Length 302;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
|
|
|
|
|
|
Db 137 AAAAPF 142

RESULT 8
AAy08994

ID AAY08994 standard; Protein; 362 AA.

XX
AC AAY08994;

DT 20-AUG-1999 (first entry)

DE Rice NAD+ isocitrate dehydrogenase protein.

KW NAD+ isocitrate dehydrogenase; corn; rice; soybean; wheat; potato;
KW tobacco; inhibitor; herbicide; plant.

OS Oryza sativa.

PN WO9928479-A1.

PD 10-JUN-1999.

PF 01-DEC-1998; 98WO-US25427.

PR 02-DEC-1997; 97US-0067388.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Allen SM, Falco SC;

DR WPI; 1999-385383/32.

XX N-PSDB; AAX78203.

PT DNA encoding an isocitrate dehydrogenase, used to identify
PT isocitrate dehydrogenase inhibitors that can be used as herbicides

PS Claim 1a; Page 32-33; 42pp; English.

XX This invention describes the isolation of novel NAD+ isocitrate
CC dehydrogenase proteins and their encoding nucleic acids from corn, rice,
CC soybean and wheat. The products of the invention can be used in methods
CC to alter the level of expression of a plant isocitrate dehydrogenase in
CC a host cell. NAD+ isocitrate dehydrogenase nucleic acid can be used as a
CC probe or primer to obtain a nucleic acid fragment encoding an isocitrate
CC dehydrogenase. The isocitrate dehydrogenase protein may be used as a
CC target to facilitate design and/or identification of inhibitors that
CC can be used as herbicides.

SQ Sequence 362 AA;

Query Match 100.0%; Score 29; DB 20; Length 362;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
|
|
|
|
|
|

Db 19 AAAAPF 24

RESULT 9

ID AAB62288 standard; Protein; 362 AA.

XX AAB62288;

DT 29-JUN-2001 (first entry)

DE Rice isocitrate dehydrogenase (clone ris6.pk0004.b3).

KW Isocitrate dehydrogenase; corn; rice; soybean; wheat; transgenic;
KW carbon metabolic pathway; citric acid cycle; ammonia assimilation;
KW herbicide.

OS Oryza sativa.

XX Key Location/Qualifiers

FT Misc-difference 58 /label= unsure

FT /note= "encoded by GCG"

PN US6204039-B1.

PD 20-MAR-2001.

PF 19-NOV-1998; 98US-0196520.

PR 02-DEC-1997; 97US-0067388.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Allen SM, Falco SC;

DR WPI; 2001-280619/29.

DR N-PSDB; AAF57652.

PT New plant isocitrate dehydrogenase polynucleotides, for modulating the
PT levels of isocitrate dehydrogenase or citric acid cycle activity in
PT plants, or as genetic tools to enhance or alter carbon or nitrogen
PT metabolic pathways

PS Claim 1; Columns 25-28; 22pp; English.

XX The invention provides nucleic acid sequences encoding corn, rice,
CC soybean and wheat isocitrate dehydrogenases. The nucleic acid sequences
CC are useful for producing or altering the levels of isocitrate
CC dehydrogenase in plants. These are also useful for creating transgenic
CC plants in which the isocitrate dehydrogenase proteins are present at
CC higher or lower levels than normal, or in cell types or developmental
CC stages in which they are not normally found. They are also useful as
CC genetic tools to enhance or alter carbon or nitrogen metabolic pathways,
CC which in turn provide mechanisms for modulating the citric acid cycle
CC and ammonia assimilation in plant cells. The isocitrate dehydrogenase
CC proteins may be used as targets to facilitate design and identification
CC of inhibitors of these enzymes that may be useful as herbicides. The
CC present sequence represents the rice isocitrate dehydrogenase.

SQ Sequence 362 AA;

Query Match 100.0%; Score 29; DB 22; Length 362;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
|
|
|
|
|
|
Db 19 AAAAPF 24

RESULT 10
AAW52823

ID AAW52823 standard; Protein; 389 AA.
XX
AC AAW52823;
XX
DT 08-SEP-1998 (first entry)
XX
DE Human reciprocal TFE3/Nono fusion product from cell line UOK109.
XX
KW PRC; papillary renal cell carcinoma; TFE3; transcription factor;
fusion protein; translocation; diagnosis; treatment; Nono; p54-nrb.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Protein 1..389
FT /label= TFE3/Nono
FT /note= "reciprocal fusion protein"
FT Misc-difference 1..295
FT /label= TFE3
FT /note= "transcription factor sequence"
FT Misc-difference 296..389
FT /label= Nono
FT /note= "also known as p54-nrb"
XX
PN WO9806871-A1.
XX
PD 19-FEB-1998.
XX
PE 13-AUG-1997; 97WO-GB02209.
XX
PR 13-AUG-1996; 96GB-0016986.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX
PI Clark J, Cooper C, Shipley J;
XX
DR WPI; 1998-159557/14.
DR N-PSDB; AAV20959.
XX
PT Diagnosing papillary renal cell carcinoma by detecting gene
trans-location - resulting in fusion of TFE3 gene with some other
PT reagents, also related vectors, transformed cells, specific binding
PT sequences
XX
PS Claim 9; Fig 3B; 71pp; English.
XX
CC This sequence represents a novel reciprocal fusion protein constructed
from the Nono protein (also known as p54-nrb) and the transcription
factor TFE3 which is found in cell line UOK109 and used in a method for
the diagnosis, prophylactic and therapeutic treatment of papillary renal
cell carcinoma. The translocation t(X;1) (p11.2;q21.2) found in papillary
renal cell carcinoma (PRC) associated protein (PRC) results in a fusion
of the TFE3 gene with a new chromosome 1 gene designated PRC
(at 1q21.2), resulting in expression of a fusion protein between the
N-terminus of PRC and almost the whole of the TFE3 gene. Normal TFE3
transcripts are no longer produced. Two other fusion partners for TFE3
have also been detected; Nono, from a invX (p11.2; q13-24 or 12)
translocation and the PSF splice factor gene, resulting in t(X;1)
(p11.2;p34). These trans-locations define a subgroup of PRC generally
encountered in patients younger than 25.
XX
SQ Sequence 389 AA;

Query Match 100.0%; Score 29; DB 19; Length 389;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
| | | | |
DB 137 AAAAPF 142

RESULT 11
AAG90077
ID AAG90077 standard; Protein; 397 AA.
XX
AC AAG90077;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 3831.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
organic acid synthesis.
XX
KM Corynebacterium glutamicum.
XX
OS
XX
PN EP1108790-A2.
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
DR N-PSDB; AAH65296.
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
PS Claim 17; SEQ ID NO: 3831; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
from Coryneform bacterium, and identifying a homologue of a gene derived
from Coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from the
European Patent Office.
XX
SQ Sequence 397 AA;

Query Match 100.0%; Score 29; DB 22; Length 397;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
| | | | |
DB 45 AAAAPF 50
RESULT 12
AAB79052
ID AAB79052 standard; Protein; 397 AA.
XX
AC AAB79052;
XX
DT 30-APR-2001 (first entry)

XX DE Corynebacterium glutamicum HA protein sequence SEQ ID NO:60.
XX KW Corynebacterium glutamicum; homeostasis; adaptation; HA protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW diagnosis; Corynebacterium diptheriae; genetic engineering;
XX KW Brevibacterium; environmental condition.
OS Corynebacterium glutamicum.
XX WO200100842-A2.
XX PN
XX PD 04-JAN-2001.
XX PF 23-JUN-2000; 2000WO-IB00911.
XX PR 25-JUN-1999; 99US-0141031.
PR 08-JUL-1999; 99DE-1031636.
PR 09-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032126.
PR 09-JUL-1999; 99DE-1032127.
PR 09-JUL-1999; 99DE-1032128.
PR 09-JUL-1999; 99DE-1032129.
PR 09-JUL-1999; 99DE-1032226.
PR 14-JUL-1999; 99DE-1032920.
PR 14-JUL-1999; 99DE-1032922.
PR 14-JUL-1999; 99DE-1032924.
PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1032930.
PR 14-JUL-1999; 99DE-1032933.
PR 14-JUL-1999; 99DE-1032935.
PR 14-JUL-1999; 99DE-1032973.
PR 14-JUL-1999; 99DE-1033002.
PR 14-JUL-1999; 99DE-1033003.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041390.
PR 31-AUG-1999; 99DE-1041391.
PR 03-SEP-1999; 99DE-1042088.
XX PA (BADI) BASF AG.
XX PI Pompejus M, Kroeger B, Schroeder H, zelder O, Haberhauer G;
XX DR WPI; 2001-061974/07.
XX DR N-PSDB; AAF71167.
XX PT New isolated Corynebacterium glutamicum nucleic acid for production or
PT modulation of production of fine chemicals such as amino acids,
PT nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins
PT or enzymes -
XX PS Claim 20; Page 219-220; 712pp; English.
XX CC AAF71138 to AAF71357 encode the Corynebacterium glutamicum homeostasis
CC and adaptation (HA) proteins given in AAB79023 to AAB79242. The
CC C. glutamicum HA genes (I) can be used in vectors for expression in host
CC cells and production of fine chemicals, such as, an organic acid,
CC proteinogenic or nonproteinogenic amino acid (preferred), purine or
CC pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated
CC fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,
CC polyketide or enzyme. The amino acids produced can be lysine, glutamine,
CC glutamate, alanine, aspartate, glycine, serine, threonine, methionine,
CC cysteine, valine, leucine, isoleucine, arginine, proline, histidine,
CC tyrosine, phenylalanine, or tryptophan. The fine chemical production can
CC be modulated. The presence of (I) or HA proteins encoded by then are
CC used for diagnosing the presence or activity of Corynebacterium
CC diptheriae. (I) can be used to map the C. glutamicum genome or can be

CC used as markers for genetically engineered Corynebacterium or
CC Brevibacterium. The HA proteins encoded by the (I) are used to maintain
CC homeostasis in C. glutamicum or help the microorganism to adapt to
CC different environmental conditions.
XX SQ Sequence 397 AA;
Query Match 100.0%; Score 29; DB 22; Length 397;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
Db 45 AAAAPF 50
RESULT 13
AAM51979
ID AAM51979 standard; Protein; 405 AA.
XX AC AAM51979;
XX DT 06-FEB-2002 (first entry)
XX DE Barley biotin synthase protein #2.
XX KW Barley; biotin synthase; biotin biosynthesis; herbicide; seedling;
KW plant development; clone bsh1.pk0005.d10; transgenic plant.
XX OS Hordeum vulgare.
XX PN US2001039042-A1.
XX PD 08-NOV-2001.
XX PF 19-DEC-2000; 2000US-0740288.
XX PR 21-DEC-1999; 99US-172929P.
XX PA (ALLE/) ALLEN S M.
PA (KINN/) KINNEY A J.
PA (MIAO/) MIAO G.
PA (OROZ/) OROZCO E M.
XX PI Allen SM, Kinney AJ, Miao G, Orozco EM;
XX DR WPI; 2002-040723/05.
XX DR N-PSDB; ABA01972.
XX PT New polypeptides, useful as targets for herbicide discovery, and as
PT probes for genetic and physical mapping of genes of which they are
PT part, or creating transgenic plants, comprises biotin synthase
PT polypeptides and encoding polynucleotides -
XX PS Claim 19; Fig 1; 46pp; English.
XX CC The present invention provides the protein and coding sequences of biotin
CC synthase enzymes from barley, wheat, maize, soybean and the prickly
CC poppy. These sequences can be used to produce transgenic plants which
CC express different levels of the gene, or express it at different times in
CC plant development. They can also be used as a target in the production of
CC herbicides. The present sequence is a barley biotin synthase protein
CC obtained from clone bsh1.pk0005.d10, which was derived from a barley
CC seedling.
XX SQ Sequence 405 AA;
Query Match 100.0%; Score 29; DB 23; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
|||||

Db 11 AAAAPF 16

RESULT 14
AAB70694
ID AAB70694 standard; Protein; 410 AA.
XX
AC AAB70694;
XX
DT 17-MAY-2001 (first entry)
XX
DE Mouse DEC2a protein sequence SEQ ID NO:14.
XX
KM DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation.
XX
OS Mus musculus.
XX
PN WO200114551-A1.
XX
PD 01-MAR-2001.
XX
PF 19-JUN-2000; 2000WO-JP03991.
XX
PR 19-AUG-1999; 99JP-0233286.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Fujimoto K, Shln M, Kato Y;
XX
DR WPI; 2001-202935/20.
DR N-PSDB; AAF74777.
XX
PT DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
PT and proliferation -
XX
PS Claim 2; Page 77-80; 83pp; Japanese.
XX
CC The present invention describes a basic helix loop helix (bHLH) type
CC transcription factor designated DEC2. DEC2 can be used as a tool in
CC the development of drugs for the treatment and prevention of disorders
CC involving cell differentiation and proliferation. The present sequence
CC represents the specifically claimed mouse DEC2a protein, as given in the
CC present invention.
XX
SQ Sequence 410 AA;

QY 1 AAAAPF 6
Db 321 AAAAPF 326

Query Match 100.0%; Score 29; DB 22; Length 410;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15
AAB70692
ID AAB70692 standard; Protein; 482 AA.
XX
AC AAB70692;
XX
DT 17-MAY-2001 (first entry)
XX
DE Human DEC2a protein sequence SEQ ID NO:2.
XX
KM DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation.
XX
OS Homo sapiens.
XX
PN WO200114551-A1.

XX
PD 01-MAR-2001.
XX
PF 19-JUN-2000; 2000WO-JP03991.
XX
PR 19-AUG-1999; 99JP-0233286.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Fujimoto K, Shin M, Kato Y;
XX
DR WPI; 2001-202935/20.
DR N-PSDB; AAF74767.
XX
PT DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
PT and proliferation -
XX
PS Claim 2; Page 55-59; 83pp; Japanese.
XX
CC The present invention describes a basic helix loop helix (bHLH) type
CC transcription factor designated DEC2. DEC2 can be used as a tool in
CC the development of drugs for the treatment and prevention of disorders
CC involving cell differentiation and proliferation. The present sequence
CC represents the specifically claimed human DEC2a protein, as given in the
CC present invention.
XX
SQ Sequence 482 AA;

QY 1 AAAAPF 6
Db 380 AAAAPF 385

Query Match 100.0%; Score 29; DB 22; Length 482;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 16
AAB70693
ID AAB70693 standard; Protein; 484 AA.
XX
AC AAB70693;
XX
DT 17-MAY-2001 (first entry)
XX
DE Human DEC2b protein sequence SEQ ID NO:12.
XX
KM DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation.
XX
OS Homo sapiens.
XX
PN WO200114551-A1.
XX
PD 01-MAR-2001.
XX
PF 19-JUN-2000; 2000WO-JP03991.
XX
PR 19-AUG-1999; 99JP-0233286.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Fujimoto K, Shin M, Kato Y;
XX
DR WPI; 2001-202935/20.
DR N-PSDB; AAF74776.
XX
PT DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
PT and proliferation -
XX
PS Claim 2; Page 69-72; 83pp; Japanese.

XX The present invention describes a basic helix loop helix (bHLH) type
CC transcription factor designated DEC2. DEC2 can be used as a tool in
CC the development of drugs for the treatment and prevention of disorders
CC involving cell differentiation and proliferation. The present sequence
CC represents the specifically claimed human DEC2b protein, as given in the
CC present invention.
XX
SQ Sequence 484 AA;

Query Match 100.0%; Score 29; DB 22; Length 484;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
 |||||
DB 382 AAAAPF 387

RESULT 17
AAW52827
ID AAW52827 standard; Protein; 618 AA.
XX
AC AAW52827;
XX
DT 08-SEP-1998 (first entry)
XX
DE Human chromosome X TFE3 protein.
XX
KW PRC; papillary renal cell carcinoma; TFE3; transcription factor;
KW fusion protein; translocation; diagnosis; treatment; chromosome X.
XX
OS Homo sapiens.
XX
PN WO9806871-A1.
XX
PD 19-FEB-1998.
XX
PF 13-AUG-1997; 97WO-GB02209.
XX
PR 13-AUG-1996; 96GB-0016986.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX
PI Clark J, Cooper C, Shipley J;
XX
DR WPI: 1998-159557/14.
DR N-PSDB; AAV20963.
XX
PT Diagnosing papillary renal cell carcinoma by detecting gene
PT trans-location - resulting in fusion of TFE3 gene with some other
PT gene, also related vectors, transformed cells, specific binding
PT reagents, peptide(s) encoded by fusions and therapeutic anti-sense
PT sequences
XX
PS Claim 13; Fig 5B; 71pp; English.
XX
CC This sequence represents the human transcription factor TFE3 which is
CC used in a method for the diagnosis, prophylactic and therapeutic
CC treatment of papillary renal cell carcinoma (PRCC). The translocation
CC t(X;1) (p11.2;q21.2) found in PRCC results in a fusion of the
CC transcription factor, TFE3 gene with a new chromosome 1 gene designated
CC PRC (at 1q21.2), resulting in expression of a fusion protein between the
CC N-terminus of PRC and almost the whole of the TFE3 gene. Normal TFE3
CC transcripts are no longer produced. Two other fusion partners for TFE3
CC have also been detected; Nono, from a invX (p11.2; q13-24 or 12)
CC translocation and the PSF splice factor gene, resulting in t(X;1)
CC (p11.2;p34). These trans-locations define a subgroup of PRCC generally
CC encountered in patients younger than 25.
XX
SQ Sequence 618 AA;

Query Match 100.0%; Score 29; DB 19; Length 618;

Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
 |||||
DB 137 AAAAPF 142

RESULT 18
ABB62129
ID ABB62129 standard; Protein; 723 AA.
XX
AC ABB62129;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 13179.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
DR N-PSDB; ABL06232.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 13179; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 723 AA;

Query Match 100.0%; Score 29; DB 22; Length 723;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
 |||||
DB 150 AAAAPF 155

RESULT 19
ABB69083
ID ABB69083 standard; Protein; 747 AA.
XX
AC ABB69083;

```

XX 26-MAR-2002 (first entry)
DT
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 34041.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL13186.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 34041; 21pp + Sequence Listing; English.
PS
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
XX
SQ Sequence 747 AA;

```

Query Match 100.0%; Score 29; DB 22; Length 747;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAAAPF 6
Db 294 AAAAPF 299

RESULT 20
AAB23271
ID AAB23271 standard; Protein; 860 AA.
XX
XX AAB23271;
AC
XX
XX 02-FEB-2001 (first entry)
DT
XX
XX Balanus amphitrite adhesion/metamorphosis-related protein Bcs-5.
DE
XX
KW Adhesion/metamorphosis-related protein Bcs-5; barnacle; larva-specific;
KW adhesion inhibition; metamorphosis inhibition; compound screening;
KW antifouling composition.
XX
XX Balanus amphitrite.
OS
XX
XX JP2000228985-A.
PN
XX

```

```

PD 22-AUG-2000.
XX
XX 09-FEB-1999; 99JP-0031067.
PF
XX 09-FEB-1999; 99JP-0031067.
PR
XX (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX
XX WPI; 2000-649634/63.
DR N-PSDB; AAA97605.
XX
XX Novel barnacle kipsis larva-specific adhesion/metamorphosis-related
PT gene useful for identifying inhibitors of barnacle adhesion -
PT
XX
XX Claim 5; Page 24-28; 32pp; Japanese.
PS
XX
XX The invention relates to six larva-specific adhesion/metamorphosis-
CC related genes from the barnacle Balanus amphitrite (CDNAs given in
CC AAA97601-A97606) and to the proteins they encode (AAB23267-B23272).
CC The genes and the proteins can be used for screening for a
CC substance that inhibits the adhesion or metamorphosis of barnacle
CC larvae, which may be useful in antifouling compositions for use in
CC the shipping industry. The present sequence represents the Balanus
CC amphitrite adhesion/metamorphosis-related protein Bcs-5.
CC
XX
SQ Sequence 860 AA;

```

Query Match 100.0%; Score 29; DB 21; Length 860;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAAAPF 6
Db 319 AAAAPF 324

```

```

RESULT 21
AAW43821
ID AAW43821 standard; peptide; 9 AA.
XX
XX AAW43821;
AC
XX
XX 20-APR-1998 (first entry)
DT
XX

```

Specific human leukocyte antigen binding peptide #25.

Human leukocyte antigen; HLA; vaccine; bacterial infection; viral;
cytotoxic T-cell; CTL; immunogenic peptide; cancer; poly-alanine.

Synthetic.

```

XX
XX W09734617-A1.
PN
XX
XX 25-SEP-1997.
PD
XX
XX 21-MAR-1997; 97WO-US04451.
PE
XX 20-MAR-1997; 97US-0821739.
PR 21-MAR-1996; 96US-0013833.
XX
XX (CYTE-) CYTEL CORP.
PA
XX
XX Celis E, Grey HM, Kubo RT, Sette A;
PI
XX
XX WPI; 1997-489250/45.
DR
XX
XX Specific human leukocyte antigen binding peptide - used in vaccines
PT for the treatment and prevention of e.g. bacterial or viral
PT infection and cancer
XX
XX Claim 19; Page 37; 49pp; English.
PS
XX
XX The present sequence represents a specific example of an immunogenic
CC

```

CC peptide which was used in a new method of inducing a cytotoxic T cell
CC (CTL) response against a preselected antigen in a patient. The method
CC comprises contacting CTLs from the patient with the immunogenic peptide
CC (containing defined motifs) which binds one of the four HLA MHC products
CC HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1, the peptide having a dissociation
CC constant (Kd) of less than 5 x 10⁻⁷ M. Immunogens are viral, e.g human
CC immunodeficiency virus type I (HIV-1), hepatitis B virus (HBV) and
CC hepatitis C virus (HCV) or cancer antigens and are used in vaccines for
CC the prevention and treatment of viral infection and cancer. The
CC immunogens may be administered to the patient as a nucleic acid encoding
CC the peptide (gene vaccine).
XX
SQ Sequence 9 AA;

Query Match 89.7%; Score 26; DB 18; Length 9;
Best Local Similarity 83.3%; Pred. No. 7.8e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAPF.6
11111:
Db 4 AAAAPY 9

RESULT 22
AAW43809
ID AAW43809 standard; peptide; 9 AA.
XX
AC AAW43809;
XX
DT 20-APR-1998 (first entry)
XX
DE Specific human leukocyte antigen binding peptide #13.
XX
KW Human leukocyte antigen; HLA; vaccine; bacterial infection; viral;
KW cytotoxic T-cell; CTL; immunogenic peptide; cancer; poly-alanine.
OS Synthetic.
XX
PN WO9734617-A1.
XX
PD 25-SEP-1997.
XX
PF 21-MAR-1997; 97WO-US04451.
XX
PR 20-MAR-1997; 97US-0821739.
PR 21-MAR-1996; 96US-0013833.
XX
PA (CYTE-) CYTEL CORP.
XX
PI Celis E, Grey HM, Kubo RT, Sette A;
XX
DR WPI; 1997-489250/45.
XX
PT Specific human leukocyte antigen binding peptide - used in vaccines
PT for the treatment and prevention of e.g. bacterial or viral
PT infection and cancer
XX
PS Claim 19; Page 37; 49pp; English.
XX
CC The present sequence represents a specific example of an immunogenic
CC peptide which was used in a new method of inducing a cytotoxic T cell
CC (CTL) response against a preselected antigen in a patient. The method
CC comprises contacting CTLs from the patient with the immunogenic peptide
CC (containing defined motifs) which binds one of the four HLA MHC products
CC HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1, the peptide having a dissociation
CC constant (Kd) of less than 5 x 10⁻⁷ M. Immunogens are viral, e.g human
CC immunodeficiency virus type I (HIV-1), hepatitis B virus (HBV) and
CC hepatitis C virus (HCV) or cancer antigens and are used in vaccines for
CC the prevention and treatment of viral infection and cancer. The
CC immunogens may be administered to the patient as a nucleic acid encoding
CC the peptide (gene vaccine).
XX
SQ Sequence 9 AA;

Query Match 89.7%; Score 26; DB 18; Length 9;
Best Local Similarity 83.3%; Pred. No. 7.8e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAPF.6
11111:
Db 4 AAAAPY 9

RESULT 23
AAU99695
ID AAU99695 standard; Peptide; 14 AA.
XX
AC AAU99695;
XX
DT 07-OCT-2002 (first entry)
XX
DE Peptide encoded by vectors PENTR/D-TOPO and pCDNAGW-DTOPO.
XX
KW Topoisomerase recognition site; topoisomerase; pCDNAGWDT(sc);
KW PENTR-DT(sc); pCDNA-DEST41; PENTR/D-TOPO; PENTR/SD/D-TOPO;
KW pCDNA3.2/V5/GWD-TOPO; pCDNA6.2/V5/GWD-TOPO; recombinational cloning;
KW gene targeting; mutation; vector.
XX
OS Synthetic.
XX
PN WO200246372-A1.
XX
PD 13-JUN-2002.
XX
PF 07-DEC-2001; 2001WO-US45773.
XX
PR 08-DEC-2000; 2000US-254510P.
PR 11-DEC-2000; 2000US-0732914.
PR 14-SEP-2001; 2001US-318902P.
PR 28-SEP-2001; 2001US-326092P.
PR 27-NOV-2001; 2001US-333124P.
XX
PA (INVT-) INVITROGEN CORP.
XX
PI Chesnut JD, Carrino J, Leong L, Madden K, Gleeson M, Fan J;
PI Brasch MA, Cheo D, Hartley JL, Byrd DRN, Temple GF;
XX
DR WPI; 2002-519662/55.
XX
PT New isolated nucleic acid molecule comprises one or more recombination
PT sites and one or more topoisomerase recognition sites and/or one or
PT more topoisomerases, useful in recombinational cloning
XX
PS Example 3; Fig 26; 324pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I)
CC comprising: (a) one or more recombination sites; and (b) one or more
CC topoisomerase recognition sites and/or one or more topoisomerases.
CC Also included are a vector comprising the nucleic acid, a vector chosen
CC from pCDNAGWDT(sc), PENTR-DT(sc), pCDNA-DEST41, PENTR/D-TOPO,
CC PENTR/SD/D-TOPO, pCDNA3, 2/V5/GWD-TOPO or pCDNA6.2/V5/GWD-TOPO,
CC a host cell comprising the nucleic acid or vectors and an in vitro method
CC of cloning a nucleic acid molecule involving: (a) obtaining a first
CC nucleic acid molecule to be cloned; (b) mixing the first nucleic acid
CC molecule to be cloned in vitro with a second nucleic acid molecule
CC comprising at least a first topoisomerase recognition site flanked by at
CC least a first recombination site, and at least a second topoisomerase
CC recognition site flanked by at least a second recombination site, where
CC the first and second recombination sites do not recombine with each
CC other, and at least one topoisomerase; and (c) incubating the mixture
CC under conditions such that the first nucleic acid molecule to be cloned
CC is inserted into the second nucleic acid molecule between the first and
CC second topoisomerase recognition sites, thereby producing a first product
CC molecule comprising the first nucleic acid molecule to be cloned between
CC the first and second recombination sites. The method is useful for

CC cloning a nucleic acid molecule. The nucleic acid (I) is useful in
CC methods for recombinational cloning and facilitates construction of gene
CC targeting nucleic acid molecules or vectors which may be used to knockout
CC or mutate a sequence or gene of interest, particularly genes or sequences
CC within a host or host cells such as animal, plant, etc. Thus the
CC nucleic acid is most preferably used for targeting or mutating a sequence
CC of gene. The present sequence is the peptide encoded by the 5' end of the
CC topoisomerase site-containing vectors pENTR/D-TOPO and pCDNAGW-DTOPO.
XX
SQ Sequence 14 AA;

Query Match 89.7%; Score 26; DB 23; Length 14;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
:|||||
Db 7 SAAAPF 12

RESULT 24
AAU53233
ID AAU53233 standard; Protein; 83 AA.
XX
AC AAU53233;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #14129.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PE 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59559.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 14428; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct-sequences.
XX
SQ Sequence 83 AA;

Query Match 89.7%; Score 26; DB 22; Length 83;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
:|||||
Db 8 AAAAPY 13

RESULT 25
AAB32830
ID AAB32830 standard; Protein; 138 AA.
XX
AC AAB32830;
XX
DT 25-JAN-2001 (first entry)
XX
DE Eucalyptus grandis transcription factor protein sequence #288.
XX
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; EREBS; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB.
XX
OS Eucalyptus grandis.
XX
PN WO200053724-A2.
XX
PD 14-SEP-2000.
XX
PE 09-MAR-2000; 2000WO-US06112.
XX
PR 11-MAR-1999; 99US-0266513.
PR 18-AUG-1999; 99US-0149485.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Wood M, McGrath A, Shenk MA, Glenn M;
XX
DR WPI; 2000-579369/54.
XX
PT New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide
PT -
XX
PS Claim 8; Page 320; 747pp; English.
XX
CC The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
CC transcription factor. The transcription factor may be used to produce a
CC plant having modified gene expression such as a woody plant e.g. a
CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
CC to modify the activity of a polypeptide in a plant. The transcription
CC factors of the present invention are members from the following families
CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
CC zipper, LIM domain, AP2 and EREBS, zinc finger domains of type 2
CC Cys2His2, CCAAT box elements and MYB.
XX

SQ Sequence 138 AA;

Query Match 89.7%; Score 26; DB 21; Length 138;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
11:111
Db 11 AASAPF 16

RESULT 26

ABG09853

ID ABG09853 standard; Protein; 159 AA.

XX AC ABG09853;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9844.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR MPI: 2001-639362/73.

DR N-PSDB; AAS74040.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID No 40212; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 159 AA;

Query Match 89.7%; Score 26; DB 22; Length 159;
Best Local Similarity 83.3%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
:11111
Db 87 SAAAPF 92

RESULT 27

AAW33276

ID AAW33276 standard; Protein; 204 AA.

XX AC AAW33276;

DT 17-FEB-1998 (first entry)

DE S. fradiae tylosin biosynthesis gene product tylJ.

XX Tylosin; biosynthetic gene product; production; antibiotic; tylJ.

OS Streptomyces fradiae.

PN US5672497-A.

PD 30-SEP-1997.

PF 21-DEC-1995; 95US-0575843.

PR 12-MAY-1989; 89US-0351350.

PR 21-MAR-1986; 86US-0842330.

PR 25-JUL-1986; 86US-0890670.

PR 24-FEB-1987; 87US-0018237.

PR 06-AUG-1991; 91US-0742222.

PR 28-JUL-1993; 93US-0107232.

PR 17-FEB-1994; 94US-0198672.

PR 21-DEC-1995; 95US-0575843.

PA (ELIL) LILLY & CO ELI.

PI Cox KL, Fishman SE, Hersherberger CL, Seno ET;

DR MPI: 1997-488860/45.

DR N-PSDB; AAT58686.

PT DNA encoding Streptomyces fradiae tylosin biosynthesis gene products
PT - for increasing tylosin production in Streptomyces spp.

PS Claim 22; Columns 21-38; 38pp; English.

CC The present sequence is the Streptomyces fradiae tylosin
CC biosynthetic gene product tylJ, useful to increase the production
CC of the antibiotic tylosin in Streptomyces spp..

SQ Sequence 204 AA;

OY 1 AAAAPF 6
11:111
Db 194 AASAPF 199

RESULT 28

AAR22377

ID AAR22377 standard; Protein; 210 AA.

XX AC AAR22377;

DT 17-AUG-1992 (first entry)

XX

DE Antigen mc-4c.
XX
KW Oocysts; sporozoite.
XX
OS Elmeria maxima.
XX
PN WO9204461-A.
XX
PD 19-MAR-1992.
XX
PF 05-SEP-1991; 91WO-US06431.
XX
PR 12-SEP-1990; 90US-0581694.
XX
PA (GENE-) GENEX CORP.
PA (FARH) HOECHST AG.
XX
PI Jacobson JW, Strausberg RL, Wilson SD, Pope SH, Strausberg SL;
PI Raether W;
XX
DR WPI: 1992-114366/14.
DR N-PSDB; AAQ23077.
XX
PT Vaccine against avian coccidiosis - comprising recombinant
PT Elmeria antigen mc-4c, mc-5c or mc-30c gene, etc., or
PT microorganisms expressing them
XX
PS Claim 2; Page 64 + Fig 2; 94pp; English.
XX
CC To identify antigens of E. maxima, expression libraries were prepd.
CC in lambda vector, lambda gtl, using cDNA prepd. from polyA mRNA
CC isolated from E. maxima oocysts. The cDNA expression library was
CC screened with rat antiserum raised against the sporozoite stage of
CC E. maxima. The library to be screened was plated on a host that
CC allows lysis and plaque formation. Following induction of the
CC antigens encoded by the phage, the plaques were transferred to
CC nitrocellulose filters. Positive phage were identified after
CC screening with the rat anti-E. maxima sporozoite antiserum.
CC The cDNA inserts from the positive clones were cloned into
CC bacteriophage M13 and subjected to sequence analysis.
CC E. maxima antigen mc-4c (21.7 kD) was identified.
XX
SQ Sequence 210 AA;

Query Match 89.7%; Score 26; DB 13; Length 210;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 180 AAAAPY 185

RESULT 29
AAG91087
ID AAG91087 standard; Protein; 325 AA.
XX
AC AAG91087;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 4841.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.

XX
PR 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI: 2001-376931/40.
DR N-PSDB; AAH66306.
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
PS Claim 17; SEQ ID NO: 4841; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 325 AA;

Query Match 89.7%; Score 26; DB 22; Length 325;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 40 AAAAPY 45

RESULT 30
AAB79630
ID AAB79630 standard; Protein; 325 AA.
XX
AC AAB79630;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum SMP protein sequence SEQ ID NO:776.
XX
KW Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;
KW diagnosis; Corynebacterium diphteriae; evolutionary study.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100844-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB00943.
XX
PR 25-JUN-1999; 99US-0141031.
PR 08-JUL-1999; 99DE-1031412.

| | | | | |
|----|---|--|-----------|--|
| PR | 08-JUL-1999; | 99DE-1031413. | RESULT 31 | |
| PR | 08-JUL-1999; | 99DE-1031419. | AAG31744 | |
| PR | 08-JUL-1999; | 99DE-1031420. | ID | AAG31744 standard; Protein; 327 AA. |
| PR | 08-JUL-1999; | 99DE-1031424. | XX | |
| PR | 08-JUL-1999; | 99DE-1031428. | AC | AAG31744; |
| PR | 08-JUL-1999; | 99DE-1031431. | XX | |
| PR | 08-JUL-1999; | 99DE-1031433. | DT | 17-OCT-2000 (first entry) |
| PR | 08-JUL-1999; | 99DE-1031434. | XX | |
| PR | 08-JUL-1999; | 99DE-1031510. | DE | Arabidopsis thaliana protein fragment SEQ ID NO: 38172. |
| PR | 08-JUL-1999; | 99DE-1031562. | XX | |
| PR | 08-JUL-1999; | 99DE-1031634. | KW | Protein identification; signal transduction pathway; |
| PR | 09-JUL-1999; | 99DE-1032180. | KW | hybridisation assay; genetic mapping; gene expression control; promoter; |
| PR | 09-JUL-1999; | 99DE-1032227. | KW | termination sequence. |
| PR | 09-JUL-1999; | 99DE-1032230. | XX | |
| PR | 09-JUL-1999; | 99US-0143208. | OS | Arabidopsis thaliana. |
| PR | 14-JUL-1999; | 99DE-1032924. | XX | |
| PR | 14-JUL-1999; | 99DE-1032973. | PN | EP1033405-A2. |
| PR | 14-JUL-1999; | 99DE-1033005. | XX | |
| PR | 27-AUG-1999; | 99DE-1040765. | PD | 06-SEP-2000. |
| PR | 31-AUG-1999; | 99US-0151572. | XX | |
| PR | 03-SEP-1999; | 99DE-1042076. | PF | 25-FEB-2000; 2000EP-0301439. |
| PR | 03-SEP-1999; | 99DE-1042079. | XX | |
| PR | 03-SEP-1999; | 99DE-1042086. | XX | |
| PR | 03-SEP-1999; | 99DE-1042087. | XX | 25-FEB-1999; 99US-0121825. |
| PR | 03-SEP-1999; | 99DE-1042088. | PR | 05-MAR-1999; 99US-0123180. |
| PR | 03-SEP-1999; | 99DE-1042095. | PR | 09-MAR-1999; 99US-0123548. |
| PR | 03-SEP-1999; | 99DE-1042123. | PR | 23-MAR-1999; 99US-0125788. |
| PR | 03-SEP-1999; | 99DE-1042125. | PR | 25-MAR-1999; 99US-0126264. |
| XX | | | PR | 29-MAR-1999; 99US-0126785. |
| PA | (BADI) BASF AG. | | PR | 01-APR-1999; 99US-0127462. |
| XX | | | PR | 06-APR-1999; 99US-0128234. |
| PI | Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G; | | PR | 08-APR-1999; 99US-0128714. |
| XX | | | PR | 16-APR-1999; 99US-0129845. |
| DR | WPI; 2001-061975/07. | | PR | 19-APR-1999; 99US-0130077. |
| DR | N-PSDB; AAF71747. | | PR | 21-APR-1999; 99US-0130079. |
| XX | | | PR | 23-APR-1999; 99US-0130510. |
| XX | | | PR | 23-APR-1999; 99US-0130891. |
| PT | New isolated Corynebacterium glutamicum nucleic acid encoding a sugar | | PR | 28-APR-1999; 99US-0131449. |
| PT | metabolism and oxidative phosphorylation protein for production or | | PR | 30-APR-1999; 99US-0132048. |
| PT | modulation of production of fine chemicals e.g. amino acids, | | PR | 30-APR-1999; 99US-0132407. |
| XX | carbohydrates or enzymes - | | PR | 04-MAY-1999; 99US-0132484. |
| PS | Claim 20; Page 1235-1236; 1246bp; English. | | PR | 05-MAY-1999; 99US-0132485. |
| XX | | | PR | 06-MAY-1999; 99US-0132486. |
| CC | AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar | | PR | 06-MAY-1999; 99US-0132487. |
| CC | metabolism and oxidative phosphorylation (SMP) proteins given in | | PR | 07-MAY-1999; 99US-0132863. |
| CC | AAB79243 to AAB 79633 which are involved in carbon metabolism and | | PR | 11-MAY-1999; 99US-0134256. |
| CC | energy production. The C. glutamicum SMP gene can be used in vectors | | PR | 14-MAY-1999; 99US-0134218. |
| CC | (II) for expression in host cells and production or modulation of | | PR | 14-MAY-1999; 99US-0134219. |
| CC | production of fine chemicals, such as, an organic acid, a proteinogenic | | PR | 14-MAY-1999; 99US-0134370. |
| CC | or nonproteinogenic amino acid (preferred), a purine or pyrimidine base, | | PR | 18-MAY-1999; 99US-0134768. |
| CC | a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty | | PR | 19-MAY-1999; 99US-0134941. |
| CC | acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a | | PR | 20-MAY-1999; 99US-0135124. |
| CC | cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins | | PR | 21-MAY-1999; 99US-0135353. |
| CC | (III) encoded by them are used for diagnosing the presence or activity of | | PR | 24-MAY-1999; 99US-0135629. |
| CC | Corynebacterium diptheriae in a subject. (I), (II), (III) or host cells | | PR | 25-MAY-1999; 99US-0136021. |
| CC | containing them are used to map genomes of organisms related to | | PR | 27-MAY-1999; 99US-0136392. |
| CC | C. glutamicum, identify and localise C. glutamicum sequences of interest, | | PR | 28-MAY-1999; 99US-0136782. |
| CC | in evolutionary studies, in determining SMP protein regions required | | PR | 01-JUN-1999; 99US-0137222. |
| CC | for function, in modulating SMP protein activity, in modulating the | | PR | 03-JUN-1999; 99US-0137528. |
| CC | metabolism of sugars, and in modulating high-energy molecule production | | PR | 04-JUN-1999; 99US-0137502. |
| CC | in a cell (i.e. ATP, NADPH). | | PR | 07-JUN-1999; 99US-0137724. |
| XX | | | PR | 08-JUN-1999; 99US-0138094. |
| SO | Sequence 325 AA; | | PR | 10-JUN-1999; 99US-0138540. |
| | | | PR | 10-JUN-1999; 99US-0138847. |
| | | | PR | 14-JUN-1999; 99US-0139119. |
| | | | PR | 16-JUN-1999; 99US-0139452. |
| | | | PR | 16-JUN-1999; 99US-0139453. |
| | | | PR | 17-JUN-1999; 99US-0139492. |
| | | | PR | 18-JUN-1999; 99US-0139454. |
| | | | PR | 18-JUN-1999; 99US-0139455. |
| | | | PR | 18-JUN-1999; 99US-0139456. |
| | | | PR | 18-JUN-1999; 99US-0139457. |
| | | | PR | 18-JUN-1999; 99US-0139458. |
| QY | 1 AAAAPF 6 | 89.7%; Score 26; DB 22; Length 325; | | |
| | : | Best Local Similarity 83.3%; Pred. No. 1,1e+03; | | |
| Db | 40 AAAAPY 45 | Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0; | | |

PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 89.7%; Score 26; DB 21; Length 327;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAPF 6
|:||||
Db 311 ASAAPF 316

RESULT 32
AAG30330
ID AAG30330 standard; Protein; 329 AA.
XX
AC AAG30330;
XX
DT 17-OCT-2000 (first entry)

| | | | |
|----|--|----|----------------------------|
| XX | Arabidopsis thaliana protein fragment SEQ ID NO: 36240. | | |
| DE | | PR | 21-JUN-1999; 99US-0139817. |
| XX | | PR | 22-JUN-1999; 99US-0139899. |
| KW | Protein identification; signal transduction pathway; metabolic pathway; | PR | 23-JUN-1999; 99US-0140353. |
| KW | hybridisation assay; genetic mapping; gene expression control; promoter; | PR | 23-JUN-1999; 99US-0140354. |
| KW | termination sequence. | PR | 24-JUN-1999; 99US-0140695. |
| XX | | PR | 28-JUN-1999; 99US-0140823. |
| OS | Arabidopsis thaliana. | PR | 29-JUN-1999; 99US-0140991. |
| XX | | PR | 30-JUN-1999; 99US-0141287. |
| PN | EP1033405-A2. | PR | 01-JUL-1999; 99US-0141842. |
| XX | | PR | 01-JUL-1999; 99US-0142154. |
| XX | | PR | 02-JUL-1999; 99US-0142055. |
| PD | 06-SEP-2000. | PR | 06-JUL-1999; 99US-0142390. |
| XX | | PR | 08-JUL-1999; 99US-0142803. |
| PF | 25-FEB-2000; 2000EP-0301439. | PR | 09-JUL-1999; 99US-0142920. |
| XX | | PR | 12-JUL-1999; 99US-0142977. |
| XX | | PR | 13-JUL-1999; 99US-0143542. |
| PR | 25-FEB-1999; 99US-0121825. | PR | 14-JUL-1999; 99US-0143624. |
| PR | 05-MAR-1999; 99US-0123180. | PR | 15-JUL-1999; 99US-0144005. |
| PR | 09-MAR-1999; 99US-0123548. | PR | 16-JUL-1999; 99US-0144085. |
| PR | 23-MAR-1999; 99US-0125788. | PR | 16-JUL-1999; 99US-0144086. |
| PR | 25-MAR-1999; 99US-0126264. | PR | 19-JUL-1999; 99US-0144325. |
| PR | 29-MAR-1999; 99US-0126785. | PR | 19-JUL-1999; 99US-0144331. |
| PR | 01-APR-1999; 99US-0127462. | PR | 19-JUL-1999; 99US-0144332. |
| PR | 06-APR-1999; 99US-0128234. | PR | 19-JUL-1999; 99US-0144333. |
| PR | 08-APR-1999; 99US-0128714. | PR | 19-JUL-1999; 99US-0144334. |
| PR | 16-APR-1999; 99US-0129845. | PR | 19-JUL-1999; 99US-0144335. |
| PR | 19-APR-1999; 99US-0130077. | PR | 20-JUL-1999; 99US-0144352. |
| PR | 21-APR-1999; 99US-0130449. | PR | 20-JUL-1999; 99US-0144632. |
| PR | 23-APR-1999; 99US-0130510. | PR | 20-JUL-1999; 99US-0144884. |
| PR | 23-APR-1999; 99US-0130891. | PR | 21-JUL-1999; 99US-0144814. |
| PR | 28-APR-1999; 99US-0131449. | PR | 21-JUL-1999; 99US-0145086. |
| PR | 30-APR-1999; 99US-0132048. | PR | 21-JUL-1999; 99US-0145088. |
| PR | 30-APR-1999; 99US-0132407. | PR | 22-JUL-1999; 99US-0145085. |
| PR | 04-MAY-1999; 99US-0132484. | PR | 22-JUL-1999; 99US-0145087. |
| PR | 05-MAY-1999; 99US-0132485. | PR | 22-JUL-1999; 99US-0145089. |
| PR | 06-MAY-1999; 99US-0132486. | PR | 22-JUL-1999; 99US-0145192. |
| PR | 06-MAY-1999; 99US-0132487. | PR | 23-JUL-1999; 99US-0145218. |
| PR | 07-MAY-1999; 99US-0132863. | PR | 23-JUL-1999; 99US-0145224. |
| PR | 11-MAY-1999; 99US-0134256. | PR | 23-JUL-1999; 99US-0145276. |
| PR | 14-MAY-1999; 99US-0134218. | PR | 26-JUL-1999; 99US-0145276. |
| PR | 14-MAY-1999; 99US-0134219. | PR | 27-JUL-1999; 99US-0145913. |
| PR | 14-MAY-1999; 99US-0134221. | PR | 27-JUL-1999; 99US-0145918. |
| PR | 14-MAY-1999; 99US-0134370. | PR | 27-JUL-1999; 99US-0145919. |
| PR | 18-MAY-1999; 99US-0134768. | PR | 28-JUL-1999; 99US-0145951. |
| PR | 19-MAY-1999; 99US-0134941. | PR | 02-AUG-1999; 99US-0146386. |
| PR | 20-MAY-1999; 99US-0135124. | PR | 02-AUG-1999; 99US-0146388. |
| PR | 21-MAY-1999; 99US-0135353. | PR | 02-AUG-1999; 99US-0146389. |
| PR | 24-MAY-1999; 99US-0135629. | PR | 03-AUG-1999; 99US-0147038. |
| PR | 25-MAY-1999; 99US-0136021. | PR | 04-AUG-1999; 99US-0147204. |
| PR | 27-MAY-1999; 99US-0136392. | PR | 04-AUG-1999; 99US-0147302. |
| PR | 28-MAY-1999; 99US-0136782. | PR | 05-AUG-1999; 99US-0147192. |
| PR | 01-JUN-1999; 99US-0137222. | PR | 05-AUG-1999; 99US-0147260. |
| PR | 03-JUN-1999; 99US-0137528. | PR | 06-AUG-1999; 99US-0147303. |
| PR | 04-JUN-1999; 99US-0137502. | PR | 06-AUG-1999; 99US-0147416. |
| PR | 07-JUN-1999; 99US-0137724. | PR | 09-AUG-1999; 99US-0147493. |
| PR | 08-JUN-1999; 99US-0138094. | PR | 10-AUG-1999; 99US-0148171. |
| PR | 10-JUN-1999; 99US-0138540. | PR | 11-AUG-1999; 99US-0148319. |
| PR | 10-JUN-1999; 99US-0138847. | PR | 12-AUG-1999; 99US-0148341. |
| PR | 14-JUN-1999; 99US-0139119. | PR | 13-AUG-1999; 99US-0148565. |
| PR | 16-JUN-1999; 99US-0139452. | PR | 13-AUG-1999; 99US-0148684. |
| PR | 16-JUN-1999; 99US-0139453. | PR | 16-AUG-1999; 99US-0149368. |
| PR | 17-JUN-1999; 99US-0139492. | PR | 17-AUG-1999; 99US-0149175. |
| PR | 18-JUN-1999; 99US-0139454. | PR | 18-AUG-1999; 99US-0149426. |
| PR | 18-JUN-1999; 99US-0139455. | PR | 20-AUG-1999; 99US-0149722. |
| PR | 18-JUN-1999; 99US-0139456. | PR | 20-AUG-1999; 99US-0149723. |
| PR | 18-JUN-1999; 99US-0139457. | PR | 20-AUG-1999; 99US-0149929. |
| PR | 18-JUN-1999; 99US-0139458. | PR | 23-AUG-1999; 99US-0149902. |
| PR | 18-JUN-1999; 99US-0139459. | PR | 23-AUG-1999; 99US-0149930. |
| PR | 18-JUN-1999; 99US-0139461. | PR | 25-AUG-1999; 99US-0150566. |
| PR | 18-JUN-1999; 99US-0139462. | PR | 26-AUG-1999; 99US-0150884. |
| PR | 18-JUN-1999; 99US-0139463. | PR | 27-AUG-1999; 99US-0151065. |
| PR | 18-JUN-1999; 99US-0139750. | PR | 27-AUG-1999; 99US-0151066. |
| PR | 18-JUN-1999; 99US-0139763. | | |

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 89.7%; Score 26; DB 21; Length 329;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
||:||||
Db 82 AASAPF 87

RESULT 33
ABG12782
ID ABG12782 standard; Protein; 331 AA.
XX
AC ABG12782;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #12773.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens .

XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS76969.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID No 43141; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 331 AA;

Query Match 89.7%; Score 26; DB 22; Length 331;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
||:||||
Db 275 AASAPF 280

RESULT 34
AAU10957
ID AAU10957 standard; Protein; 407 AA.
XX
AC AAU10957;
XX
DT 12-MAR-2002 (first entry)
XX
DE Even-skipped homeo box 1 (EVX1).
XX
KW Even-skipped homeo box 1; EVX1; neurological disease; drug screening;
KW haplotyping; single nucleotide polymorphism; SNP; human.
XX
OS Homo sapiens.
XX
PN WO200190120-A2.

XX 29-NOV-2001.
PD
XX
XX 21-MAY-2001; 2001WO-US16559.
PF
XX 19-MAY-2000; 2000US-205437P.
PR
XX (GENA-) GENAISSANCE PHARM INC.
PA
XX Duda A, Klieem SE, Kumar AM;
PI
XX WPI: 2002-089913/12.
DR
DR N-PSDB; AAS18559.
XX
PT Novel genetic variants of even-skipped homeo box 1, EVX1 gene useful
PT for therapeutic purposes and for expressing EVX1 protein useful in
PT identifying drugs to treat neurological diseases
XX
XX Claim 28; Fig 3; 69pp; English.
PS
XX The invention relates to an isolated polynucleotide (I), comprising a
CC nucleotide sequence which is a polymorphic variant of a reference
CC sequence for the even-skipped homeo box 1 (homologue of Drosophila)
CC (EVX1) gene or its fragment, or a polymorphic variant of a reference
CC sequence for a EVX1 cDNA or its fragment. EVX1 polypeptide (II) is
CC useful for screening for drugs targeting the polypeptide, by contacting
CC the EVX1 polymorphic variant with a candidate agent and assaying for
CC binding activity. A method is described for identifying an association
CC between a trait such as a clinical response to a drug targeting EVX1 and
CC a haplotype or haplotype pair of EVX1 gene. The methods are useful
CC in developing diagnostic tests and therapeutic treatments for
CC neurological diseases. (I) is useful for studying the expression and
CC function of EVX1 and expressing EVX1 protein for use in screening for
CC candidate drugs to treat diseases related to EVX1 activity. The
CC polymorphism and haplotype data are useful for validating whether EVX1 is
CC a suitable target for drugs to treat neurological diseases, screening for
CC such drugs and reducing bias in clinical trials of such drugs. (I) is
CC useful for therapeutic purposes. (I) is useful for determining if an
CC individual has one of the haplotypes 1-4 or the haplotype pairs.
CC Establishing the EVX1 haplotype or haplotype pair of an individual is
CC useful for improving the efficiency and reliability of several steps in
CC the discovery and development of drugs for treating diseases associated
CC with EVX1 activity e.g. neurological diseases. The haplotyping method
CC is useful to validate EVX1 as a candidate target for treating a
CC specific condition or disease predicted to be associated with EVX1
CC activity. (I) is useful for studying expression of the EVX1 isogenes
CC in vivo, for in vivo screening and testing of drugs against EVX1
CC protein and for testing the efficacy of therapeutic agents and
CC compounds for neurological diseases in a biological system. Antibody
CC raised against (II) is useful for diagnostic and prognostic formats and
CC therapeutic methods, for immunoprecipitating (II) from solution, for
CC detecting EVX1 protein isoforms in biological samples, frozen tissue
CC sections, cells which have been fixed or unfixed and prepared on slides,
CC for use in immunocytochemical, immunohistochemical and immunofluorescence
CC techniques. The present sequence represents the amino acid sequence of
CC human even-skipped homeo box 1 (EVX1).
XX
SQ Sequence 407 AA;
Query Match 89.7%; Score 26; DB 23; Length 407;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
| | | : | |
Db 295 AAASPF 300
RESULT 35
AAU10958
ID AAU10958 standard; Protein: 407 AA.
XX
AC AAU10958;

XX 12-MAR-2002 (first entry)
DT
XX Even-skipped homeo box 1 (EVX1) isoform #1.
DE
XX Even-skipped homeo box 1; EVX1; neurological disease; drug screening;
KW haplotyping; single nucleotide polymorphism; SNP; human.
KM
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 82 /note= "wild type Ala substituted by Val"
FT
FT
XX
XX WO200190120-A2.
PN
XX 29-NOV-2001.
PD
XX 21-MAY-2001; 2001WO-US16559.
PE
XX 19-MAY-2000; 2000US-205437P.
PR
XX (GENA-) GENAISSANCE PHARM INC.
PA
PI Duda A, Klieem SE, Kumar AM;
PI
XX WPI: 2002-089913/12.
DR
XX Novel genetic variants of even-skipped homeo box 1, EVX1 gene useful
PT for therapeutic purposes and for expressing EVX1 protein useful in
PT identifying drugs to treat neurological diseases
XX
XX Disclosure; Page -: 69pp; English.
PS
XX The invention relates to an isolated polynucleotide (I), comprising a
CC nucleotide sequence which is a polymorphic variant of a reference
CC sequence for the even-skipped homeo box 1 (homologue of Drosophila)
CC (EVX1) gene or its fragment, or a polymorphic variant of a reference
CC sequence for a EVX1 cDNA or its fragment. EVX1 polypeptide (II) is
CC useful for screening for drugs targeting the polypeptide, by contacting
CC the EVX1 polymorphic variant with a candidate agent and assaying for
CC binding activity. A method is described for identifying an association
CC between a trait such as a clinical response to a drug targeting EVX1 and
CC a haplotype or haplotype pair of EVX1 gene. The methods are useful
CC in developing diagnostic tests and therapeutic treatments for
CC neurological diseases. (I) is useful for studying the expression and
CC function of EVX1 and expressing EVX1 protein for use in screening for
CC candidate drugs to treat diseases related to EVX1 activity. The
CC polymorphism and haplotype data are useful for validating whether EVX1 is
CC a suitable target for drugs to treat neurological diseases, screening for
CC such drugs and reducing bias in clinical trials of such drugs. (I) is
CC useful for therapeutic purposes. (I) is useful for determining if an
CC individual has one of the haplotypes 1-4 or the haplotype pairs.
CC Establishing the EVX1 haplotype or haplotype pair of an individual is
CC useful for improving the efficiency and reliability of several steps in
CC the discovery and development of drugs for treating diseases associated
CC with EVX1 activity e.g. neurological diseases. The haplotyping method
CC is useful to validate EVX1 as a candidate target for treating a
CC specific condition or disease predicted to be associated with EVX1
CC activity. (I) is useful for studying expression of the EVX1 isogenes
CC in vivo, for in vivo screening and testing of drugs against EVX1
CC protein and for testing the efficacy of therapeutic agents and
CC compounds for neurological diseases in a biological system. Antibody
CC raised against (II) is useful for diagnostic and prognostic formats and
CC therapeutic methods, for immunoprecipitating (II) from solution, for
CC detecting EVX1 protein isoforms in biological samples, frozen tissue
CC sections, cells which have been fixed or unfixed and prepared on slides,
CC for use in immunocytochemical, immunohistochemical and immunofluorescence
CC techniques. The present sequence represents the amino acid sequence of
CC human even-skipped homeo box 1 (EVX1) isoform #1.
CC Note: The present sequence is not shown in the specification but is
CC derived from the wild type human EVX1 sequence given in figure 3 (see
CC AAU10957).

XX Sequence 407 AA;
SQ Query Match 89.7%; Score 26; DB 23; Length 407;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
111:11
Db 295 AAASPF 300
RESULT 36
AAU10959
ID AAU10959 standard; Protein; 407 AA.
AC AAU10959;
XX 12-MAR-2002 (first entry)
DT Even-skipped homeo box 1 (EVX1) isoform #2.
XX
DE Even-skipped homeo box 1; EVX1; neurological disease; drug screening;
XX
KW haplotyping; single nucleotide polymorphism; SNP; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 320
FT /note= "Wild type Pro substituted by Ala"
XX
PN WO200190120-A2.
XX
PD 29-NOV-2001.
XX
PF 21-MAY-2001; 2001WO-US16559.
XX
PR 19-MAY-2000; 2000US-205437P.
XX
PA (GENA-) GENAISSANCE PHARM INC.
XX
PI Duda A, Kilem SE, Kumar AM;
XX
DR WPI; 2002-089913/12.
XX
PT Novel genetic variants of even-skipped homeo box 1, EVX1 gene useful
PT for therapeutic purposes and for expressing EVX1 protein useful in
PT identifying drugs to treat neurological diseases
XX
PS Disclosure; Page -, 69pp; English.
XX
CC The invention relates to an isolated polynucleotide (I), comprising a
CC nucleotide sequence which is a polymorphic variant of a reference
CC sequence for the even-skipped homeo box 1 (homologue of Drosophila)
CC (EVX1) gene or its fragment, or a polymorphic variant of a reference
CC sequence for a EVX1 cDNA or its fragment. EVX1 polypeptide (II) is
CC useful for screening for drugs targeting the polypeptide, by contacting
CC the EVX1 polymorphic variant with a candidate agent and assaying for
CC binding activity. A method is described for identifying an association
CC between a trait such as a clinical response to a drug targeting EVX1 and
CC a haplotype or haplotype pair of EVX1 gene. The methods are useful
CC in developing diagnostic tests and therapeutic treatments for
CC neurological diseases. (I) is useful for studying the expression and
CC function of EVX1 and expressing EVX1 protein for use in screening for
CC candidate drugs to treat diseases related to EVX1 activity. The
CC polymorphism and haplotype data are useful for validating whether EVX1 is
CC a suitable target for drugs to treat neurological diseases, screening for
CC such drugs and reducing bias in clinical trials of such drugs. (I) is
CC useful for therapeutic purposes. (I) is useful for determining if an
CC individual has one of the haplotypes 1-4 or the haplotype pairs.
CC Establishing the EVX1 haplotype or haplotype pair of an individual is
CC useful for improving the efficiency and reliability of several steps in
CC the discovery and development of drugs for treating diseases associated

CC with EVX1 activity e.g. neurological diseases. The haplotyping method
CC is useful to validate EVX1 as a candidate target for treating a
CC specific condition or disease predicted to be associated with EVX1
CC activity. (I) is useful for studying expression of the EVX1 isogenes
CC in vivo, for in vivo screening and testing of drugs against EVX1
CC protein and for testing the efficacy of therapeutic agents and
CC compounds for neurological diseases in a biological system. Antibody
CC raised against (II) is useful for diagnostic and prognostic formats and
CC therapeutic methods, for immunoprecipitating (II) from solution, for
CC detecting EVX1 protein isoforms in biological samples, frozen tissue
CC sections, cells which have been fixed or unfixed and prepared on slides,
CC for use in immunocytochemical, immunohistochemical and immunofluorescence
CC techniques. The present sequence represents the amino acid sequence of
CC human even-skipped homeo box 1 (EVX1) isoform #2.
CC Note: The present sequence is not shown in the specification but is
CC derived from the wild type human EVX1 sequence given in figure 3 (see
CC AAU10957).
XX
SQ Sequence 407 AA;
QY Query Match 89.7%; Score 26; DB 23; Length 407;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
111:11
Db 295 AAASPF 300
RESULT 37
AAG78989
ID AAG78989 standard; Protein; 414 AA.
XX
AC AAG78989;
XX
DT 18-JAN-2002 (first entry)
XX
DE Escherichia coli LOLE protein, involved in lipoprotein localisation.
XX
KW LOLE; lipoprotein localisation; Gram-negative bacterium.
XX
OS Escherichia coli.
XX
PN JP2001231570-A.
XX
PD 28-AUG-2001.
XX
PF 24-FEB-2000; 2000JP-0047702.
XX
PR 24-FEB-2000; 2000JP-0047702.
XX
PA (MITA) MITSUI CHEM INC.
XX
DR WPI; 2001-650996/75.
DR N-PSDB; AA171758.
XX
PT New protein for controlling localization of a lipoprotein in the cell
PT of a Gram-negative microbe and a DNA encoding it
XX
PS Disclosure; Page 10-12; 17pp; Japanese.
XX
CC The present invention relates to novel proteins from Escherichia coli,
CC which control the localisation of lipoprotein in Gram-negative microbes.
CC The present sequence is one such protein, LOLE.
XX
SQ Sequence 414 AA;
QY Query Match 89.7%; Score 26; DB 22; Length 414;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
11111:

Db 92 AAAAPY 97

RESULT 38
ABB92120
ID ABB92120 standard; Protein; 416 AA.
XX
XX ABB92120;
AC
XX
XX 31-MAY-2002 (first entry)
DT
XX Herbicidally active polypeptide SEQ ID NO 1331.
DE
XX Herbicidal; plant; agriculture; herbicide.
KW
XX Arabidopsis thaliana.
OS
XX WO200210210-A2.
PN
XX 07-FEB-2002.
PD
XX 28-AUG-2001; 2001WO-EP09892.
PF
XX 28-AUG-2001; 2001WO-EP09892.
PR
XX (FARB) BAYER AG.
XX
XX Tietjen K, Weidler M;
PI
XX WPI; 2002-269010/31.
DR
XX
XX PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
XX PS Claim 5; SEQ ID NO 1331; 261pp + Sequence Listing; English.
XX
XX CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
XX SQ Sequence 416 AA;

Query Match 89.7%; Score 26; DB 23; Length 416;
Best Local Similarity 83.3%; Pred. NO. 1.5e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
I:||||
Db 400 ASAAPF 405

RESULT 39
AAB19410
ID AAB19410 standard; Protein; 431 AA.
XX
XX AAB19410;
AC
XX 06-MAR-2001 (first entry)
DT
XX An aromatic class prenyltransferase designated ATPt4.
DE
XX
XX Prenyltransferase; ATPt1; ATPt2; ATPt3; ATPt4; ATPt5; ATPt6; ATPt7;
KW ATPt8; ATPt9; ATPt10; ATPt11; ATPt12; tocopherol; homogenetic acid;
KW phytolpyrophosphate; 2-methyl-6phytylbenzoquinol; antioxidant;
KW nutritional supplement.

XX Arabidopsis sp.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 8 /note= "encoded by TCT"
FT Misc-difference 11 /note= "encoded by TTA"
FT
XX
XX WO200063391-A2.
PN
XX 26-OCT-2000.
PD
XX 14-APR-2000; 2000WO-US10368.
PF
XX 15-APR-1999; 99US-0129899.
PR 30-JUL-1999; 99US-0146461.
XX
XX (CALJ) CALGENE LLC.
PA
XX
XX Savidge B, Lassner MW, Weiss JD, Post-Beltemmiller D;
PI
XX WPI: 2000-647519/62.
DR N-PSDB; AAC61696.
DR
XX
XX An isolated nucleic acid sequence encoding prenyltransferase used to
PT transform plant cells to increase the production of tocopherols -
PT
XX
PS Claim 6; Fig 1; 114pp; English.

CC The present sequence represents a prenyltransferase. The specification
CC describes prenyltransferases designated ATPt1, ATPt2, ATPt3, ATPt4,
CC ATPt5, ATPt6, ATPt7, ATPt8, ATPt9, ATPt10, ATPt11, and ATPt12. The
CC biosynthesis of alpha-tocopherol in higher plants involves the
CC condensation of homogenetic acid and phytolpyrophosphate to form
CC 2-methyl-6phytylbenzoquinol, which can form various tocopherols. The
CC prenyltransferase polynucleotides are useful in transforming host cells
CC to alter the expression of prenyltransferase in these cells. The
CC transformed cells are used in the production of tocopherols which are
CC of use in the pharmaceutical industry as antioxidants and also in the
CC food industry as nutritional supplements.
XX
XX SQ Sequence 431 AA;

Query Match 89.7%; Score 26; DB 21; Length 431;
Best Local Similarity 83.3%; Pred. NO. 1.5e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
I:||||
Db 415 ASAAPF 420

RESULT 40
ABB81703
ID ABB81703 standard; Protein; 431 AA.
XX
XX ABB81703;
AC
XX 16-AUG-2002 (first entry)
DT
XX Arabidopsis sp. ATPt4.
DE
XX
XX Prenyltransferase; PTS; isoprenoid; Synechocystis; Arabidopsis; corn;
KW rice; wheat; leek; canola; cotton; tomato; biosynthetic flux;
KW tocopherols; ATPt4; polytransferase.
XX
XX Arabidopsis sp.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 8 /note= "Encoded by TCT"
FT Misc-difference 10

FT /note- "Encoded by TTA"
XX WO200233060-A2.
PN
XX
XX 25-APR-2002.
PD
XX 12-OCT-2001; 2001WO-US42673.
PF
XX 14-OCT-2000; 2000US-0688071.
PR
XX (MONS) MONSANTO TECHNOLOGY LLC.
PA
XX
PI Lassner MW, Savidge B, Weiss JD, Mitsky TA, Post-Beittemiller MA;
PI Valentin HE;
XX
XX WPI; 2002-463312/49.
DR N-PSDB; ABQ62166.
DR
XX
XX Novel nucleic acid sequences encoding prenyltransferase derived from
PT eukaryotic and prokaryotic sources useful for producing plants and
PT seeds with altered tocopherol content and compositions
XX
XX Example 1; Fig 2; 148pp; English.
PS
XX
XX The invention relates to a novel isolated nucleic acid sequence encoding
CC a prenyltransferase (PTS). The nucleic acid construct of the invention is
CC useful for altering isoprenoid content and for producing an isoprenoid
CC compound of interest in a host cell, such as *Synechocystis* sp., or a
CC plant cell obtained from *Arabidopsis*, corn, rice, wheat, leek, canola,
CC cotton or tomato. The construct is also useful for increasing
CC biosynthetic flux in a host cell toward production of an isoprenoid
CC compound such as tocopherol and tocotrienol. The polynucleotide is useful
CC for producing plants and plant parts, such as seeds, enriched in
CC tocopherols. The sequence shown represents *Arabidopsis* sp. ATP4, a
CC polytransferase identified in the invention.
XX
SQ Sequence 431 AA;

Query Match 89.7%; Score 26; DB 23; Length 431;
Best Local Similarity 83.3%; Pred. No. 1.5e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|:||||
Db 415 ASAAPF 420

Search completed: December 6, 2002, 13:28:03
Job time : 40 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:26:14 ; Search time 14 Seconds
(without alignments)
12.610 Million cell updates/sec

Title: US-10-033-526-4
Perfect-score: 29
Sequence: 1 AAAAPF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------------------------|
| 1 | 29 | 100.0 | 55 | 1 | US-08-366-953A-42 Sequence 42, Appl |
| 2 | 29 | 100.0 | 55 | 1 | US-08-366-953A-43 Sequence 43, Appl |
| 3 | 29 | 100.0 | 60 | 2 | US-08-484-397A-35 Sequence 35, Appl |
| 4 | 29 | 100.0 | 362 | 4 | US-09-196-520-4 Sequence 4, Appl |
| 5 | 26 | 89.7 | 4551 | 3 | US-09-320-878-1 Sequence 1, Appl |
| 6 | 26 | 89.7 | 4613 | 4 | US-09-105-537-31 Sequence 31, Appl |
| 7 | 26 | 89.7 | 11877 | 4 | US-09-105-537-6 Sequence 6, Appl |
| 8 | 25 | 86.2 | 7 | 2 | US-08-666-473-48 Sequence 48, Appl |
| 9 | 25 | 86.2 | 57 | 4 | US-08-984-151-2 Sequence 2, Appl |
| 10 | 25 | 86.2 | 64 | 3 | US-08-759-463-4 Sequence 4, Appl |
| 11 | 25 | 86.2 | 64 | 4 | US-09-328-153-4 Sequence 4, Appl |
| 12 | 25 | 86.2 | 72 | 1 | US-08-217-360-6 Sequence 6, Appl |
| 13 | 25 | 86.2 | 72 | 1 | US-08-217-360-8 Sequence 8, Appl |
| 14 | 25 | 86.2 | 72 | 1 | US-08-217-360-10 Sequence 10, Appl |
| 15 | 25 | 86.2 | 72 | 1 | US-08-217-360-12 Sequence 12, Appl |
| 16 | 25 | 86.2 | 453 | 4 | US-09-134-001C-4873 Sequence 4873, Ap |
| 17 | 25 | 86.2 | 488 | 4 | US-09-444-728-2 Sequence 2, Appl |
| 18 | 25 | 86.2 | 548 | 1 | US-08-468-700-37 Sequence 37, Appl |
| 19 | 25 | 86.2 | 548 | 1 | US-08-645-971-5 Sequence 5, Appl |
| 20 | 25 | 86.2 | 548 | 2 | US-08-468-220-35 Sequence 35, Appl |
| 21 | 25 | 86.2 | 548 | 2 | US-08-468-698-35 Sequence 35, Appl |
| 22 | 25 | 86.2 | 548 | 2 | US-08-704-706A-37 Sequence 37, Appl |
| 23 | 25 | 86.2 | 548 | 3 | US-08-890-383-6 Sequence 6, Appl |
| 24 | 25 | 86.2 | 548 | 3 | US-08-914-679A-6 Sequence 6, Appl |
| 25 | 25 | 86.2 | 548 | 4 | US-08-985-659-38 Sequence 38, Appl |
| 26 | 25 | 86.2 | 548 | 4 | US-08-194-664A-35 Sequence 35, Appl |
| 27 | 25 | 86.2 | 548 | 5 | PCT-US94-01553A-35 Sequence 35, Appl |

| | | | | | | |
|----|----|------|------|---|-------------------|-------------------|
| 28 | 25 | 86.2 | 548 | 5 | PCT-US95-10426-35 | Sequence 35, Appl |
| 29 | 25 | 86.2 | 549 | 1 | US-08-720-899-6 | Sequence 6, Appl |
| 30 | 25 | 86.2 | 549 | 1 | US-08-459-610-6 | Sequence 6, Appl |
| 31 | 25 | 86.2 | 549 | 2 | US-08-343-804-6 | Sequence 6, Appl |
| 32 | 25 | 86.2 | 549 | 2 | US-08-687-399-6 | Sequence 6, Appl |
| 33 | 25 | 86.2 | 549 | 2 | US-08-600-908A-6 | Sequence 6, Appl |
| 34 | 25 | 86.2 | 549 | 3 | US-08-683-838A-6 | Sequence 6, Appl |
| 35 | 25 | 86.2 | 549 | 4 | US-09-636-252A-6 | Sequence 6, Appl |
| 36 | 25 | 86.2 | 967 | 4 | US-09-313-677-17 | Sequence 17, Appl |
| 37 | 25 | 86.2 | 984 | 4 | US-09-171-461-14 | Sequence 14, Appl |
| 38 | 25 | 86.2 | 1706 | 2 | US-08-459-568-2 | Sequence 2, Appl |
| 39 | 25 | 86.2 | 1706 | 2 | US-08-399-411-2 | Sequence 2, Appl |
| 40 | 25 | 86.2 | 1706 | 3 | US-08-516-859A-2 | Sequence 2, Appl |
| 41 | 25 | 86.2 | 1706 | 4 | US-09-586-472-2 | Sequence 2, Appl |
| 42 | 25 | 86.2 | 1706 | 4 | US-09-528-706-2 | Sequence 2, Appl |
| 43 | 24 | 82.8 | 93 | 4 | US-08-981-392-44 | Sequence 44, Appl |
| 44 | 24 | 82.8 | 125 | 4 | US-09-357-251-12 | Sequence 12, Appl |
| 45 | 24 | 82.8 | 130 | 4 | US-08-833-876-4 | Sequence 4, Appl |

ALIGNMENTS

```
RESULT 1
US-08-366-953A-42
; Sequence 42, Application US/08366953A
; Patent No. 5766593
;
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri S.
; APPLICANT: Wright, Samuel D.
; TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 PEPTIDES
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1720
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/366,953A
; FILING DATE: 30-DEC-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook Ph.D., Robert R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-324
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 55 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-366-953A-42

Query Match      100.0%; Score 29; DB 1; Length 55;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAAAPF 6
Db      45 AAAAPF 50

RESULT 2
US-08-366-953A-43
; Sequence 43, Application US/08366953A
```


; Patent No. 5766593
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri S.
; APPLICANT: Wright, Samuel D.
; TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 PEPTIDES
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1720
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/366,953A
; FILING DATE: 30-DEC-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook Ph.D., Robert R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-324
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 55 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-366-953A-43

Query Match 100.0%; Score 29; DB 1; Length 55;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 45 AAAAPF 50

RESULT 3
US-08-484-397A-35
; Sequence 35, Application US/08484397A
; Patent No. 5869055
; GENERAL INFORMATION:
; APPLICANT: Juan, Shao-Chieh
; APPLICANT: Lichenstein, Henri S.
; APPLICANT: Wright, Samuel D.
; TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 POLYPEPTIDES
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: CA
; COUNTRY: US
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,397A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Curry, Daniel R.
; REGISTRATION NUMBER: 32,727

; REFERENCE/DOCKET NUMBER: A-324A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 805/447-1000
; TELEFAX: 805/447-1090
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-397A-35

Query Match 100.0%; Score 29; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 45 AAAAPF 50

RESULT 4
US-09-196-520-4
; Sequence 4, Application US/09196520
; Patent No. 6204039
; GENERAL INFORMATION:
; APPLICANT: Falco, Carl S.
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: Plant Isocitrate Dehydrogenase Homologs
; FILE REFERENCE: BB-1291
; CURRENT APPLICATION NUMBER: US/09/196,520
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: 60/067,388
; EARLIER FILING DATE: December 2, 1997
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (58)
; US-09-196-520-4

Query Match 100.0%; Score 29; DB 4; Length 362;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 19 AAAAPF 24

RESULT 5
US-09-320-878-1
; Sequence 1, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06

```
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4551
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-320-878-1

Query Match          89.7%; Score 26; DB 3; Length 4551;
Best Local Similarity 83.3%; Pred. No. 4.3e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 3725 AAAAPY 3730

RESULT 6
US-09-105-537-31
; Sequence 31, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-31

Query Match          89.7%; Score 26; DB 4; Length 4613;
Best Local Similarity 83.3%; Pred. No. 4.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 3787 AAAAPY 3792

RESULT 7
US-09-105-537-6
; Sequence 6, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 11877
; TYPE: PRT
```

```
; ORGANISM: Streptomyces venezuelae
US-09-105-537-6

Query Match          89.7%; Score 26; DB 4; Length 11877;
Best Local Similarity 83.3%; Pred. No. 1.1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 4123 AAAAPY 4128

RESULT 8
US-08-666-473-48
; Sequence 48, Application US/08666473
; Patent No. 5843713
; GENERAL INFORMATION:
; APPLICANT: YOSHIDA, Aruto
; APPLICANT: TAKEUCHI, Makoto
; TITLE OF INVENTION: PEPTIDE SEQUENCE THAT FORMS MUCIN SUGAR
; TITLE OF INVENTION: CHAIN AND TECHNIQUE FOR MODIFYING PROTEIN TO BE LINKED
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/666,473
; FILING DATE: 19-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP95/02238
; FILING DATE: 01-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-22101
; FILING DATE: 09-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-269111
; FILING DATE: 01-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16887/837
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-666-473-48

Query Match          86.2%; Score 25; DB 2; Length 7;
Best Local Similarity 83.3%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 1 AAATPF 6
```

```
RESULT 9
US-08-984-151-2
; Sequence 2, Application US/08984151
; Patent No. 6130366
; GENERAL INFORMATION:
; APPLICANT: HERRERA-ESTRELLA, LUIS
; APPLICANT: VAN DEN BROECK, GUIDO
; APPLICANT: VAN MONTAGU, MARC
; APPLICANT: SCHREIER, PETER
; APPLICANT: SCHELL, JOSEF
; APPLICANT: BOHNERT, HANS J.
; APPLICANT: CASHMORE, ANTONY
; APPLICANT: TIMKO, MICHAEL P.
; APPLICANT: KAUSCH, ALBERT P.
; TITLE OF INVENTION: CHIMAERIC GENE CODING FOR A TRANSIT
; TITLE OF INVENTION: PEPTIDE AND A HETEROLOGOUS POLYPEPTIDE (AS AMENDED)
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/984,151
; FILING DATE: 03-DEC-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 2121-0134P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Pisum sativum
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..57
; OTHER INFORMATION: /note="transit peptide of the
; OTHER INFORMATION: small subunit of Rubp of pea cells."
US-08-984-151-2
Query Match      86.2%; Score 25; DB 4; Length 57;
Best Local Similarity 83.3%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAAPF 6
      || |||
Db      22 AAVAPF 27

RESULT 10
US-08-759-463-4
; Sequence 4, Application US/08759463
; Patent No. 6011198
; GENERAL INFORMATION:
; APPLICANT: Kenton KO
; APPLICANT: zdenka W. KO
```

```
APPLICANT: Carlos A. Labate
APPLICANT: Antonio Graneli
; TITLE OF INVENTION: Gene Constructs and Methods for Enhancing
; TITLE OF INVENTION: Protein Levels in Photosynthetic Organisms
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/759,463
; FILING DATE: 05-DEC-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08\568,168
; FILING DATE: 06-DEC-1995
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-759-463-4
```

```
Query Match      86.2%; Score 25; DB 3; Length 64;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAAPF 6
      || |||
Db      22 AAVAPF 27
```

```
RESULT 11
US-09-328-153-4
; Sequence 4, Application US/09328153
; Patent No. 6239332
; GENERAL INFORMATION:
; APPLICANT: KO, Kenton
; APPLICANT: Labate, Carlos A.
; APPLICANT: Graneli, Antonio
; TITLE OF INVENTION: CONSTRUCTS AND METHODS FOR ENHANCING
; TITLE OF INVENTION: PROTEIN LEVELS IN PHOTOSYNTHETIC ORGANISMS
; FILE REFERENCE: 2159.1002016
; CURRENT APPLICATION NUMBER: US/09/328,153
; CURRENT FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: US 08/759,463
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 08/568,168
; PRIOR FILING DATE: 1995-12-06
; PRIOR APPLICATION NUMBER: PCT/CA96/00807
; PRIOR FILING DATE: 1996-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Pisum sativum
US-09-328-153-4
```

```
Query Match      86.2%; Score 25; DB 4; Length 64;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

OY 1 AAAAPF 6
11111
Db 22 AAVAPF 27

RESULT 12
US-08-217-360-6

; Sequence 6, Application US/08217360
; Patent No. 5530191
; GENERAL INFORMATION:
; APPLICANT: MALIGA, Pal
; TITLE OF INVENTION: METHODS FOR PRODUCING CYTOPLASMIC
; TITLE OF INVENTION: MALE STERILITY IN PLANTS AND USE THEREOF IN PRODUCTION OF HYBR
; TITLE OF INVENTION: SEED
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman,
; STREET: 1601 Market Street, Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/217,360
; FILING DATE: 24-MAR-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, Janet E.
; REGISTRATION NUMBER: 36,252
; REFERENCE/DOCKET NUMBER: Rutgers University
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-217-360-6

Query Match

Best Local Similarity 86.2%; Score 25; DB 1; Length 72;
Matches 5; Conservativity 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAAAPF 6
11111
Db 22 AAVAPF 27

RESULT 13
US-08-217-360-8

; Sequence 8, Application US/08217360
; Patent No. 5530191
; GENERAL INFORMATION:
; APPLICANT: MALIGA, Pal
; TITLE OF INVENTION: METHODS FOR PRODUCING CYTOPLASMIC
; TITLE OF INVENTION: MALE STERILITY IN PLANTS AND USE THEREOF IN PRODUCTION OF HYBR
; TITLE OF INVENTION: SEED
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman,
; STREET: 1601 Market Street, Suite 720
; CITY: Philadelphia

; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/217,360
; FILING DATE: 24-MAR-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, Janet E.
; REGISTRATION NUMBER: 36,252
; REFERENCE/DOCKET NUMBER: Rutgers University
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-217-360-8

Query Match 86.2%; Score 25; DB 1; Length 72;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservativity 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAAAPF 6
11111
Db 22 AAVAPF 27

RESULT 14
US-08-217-360-10

; Sequence 10, Application US/08217360
; Patent No. 5530191
; GENERAL INFORMATION:
; APPLICANT: MALIGA, Pal
; TITLE OF INVENTION: METHODS FOR PRODUCING CYTOPLASMIC
; TITLE OF INVENTION: MALE STERILITY IN PLANTS AND USE THEREOF IN PRODUCTION OF H
; TITLE OF INVENTION: SEED
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman,
; STREET: 1601 Market Street, Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/217,360
; FILING DATE: 24-MAR-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, Janet E.
; REGISTRATION NUMBER: 36,252
; REFERENCE/DOCKET NUMBER: Rutgers University
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044

; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-217-360-10

Query Match 86.2%; Score 25; DB 1; Length 72;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAAAPF 6
11111
Db 22 AAVAPF 27

RESULT 15
US-08-217-360-12
; Sequence 12, Application US/08217360
; Patent No. 5530191
; GENERAL INFORMATION:

; APPLICANT: MALIGA, Pal
; TITLE OF INVENTION: METHODS FOR PRODUCING CYTOPLASMIC
; TITLE OF INVENTION: MALE STERILITY IN PLANTS AND USE THEREOF IN PRODUCTION OF HYB
; TITLE OF INVENTION: SEED
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman,
; ADDRESSEE: P.C.
; STREET: 1601 Market Street, Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/217,360
; FILING DATE: 24-MAR-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, Janet E.
; REGISTRATION NUMBER: 36,252
; REFERENCE/DOCKET NUMBER: Rutgers University
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-217-360-12

Query Match 86.2%; Score 25; DB 1; Length 72;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAAAPF 6
11111
Db 22 AAVAPF 27

RESULT 16
US-09-134-001C-4873
; Sequence 4873, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4873
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4873

Query Match 86.2%; Score 25; DB 4; Length 453;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAAAPF 6
11111
Db 145 AATAPF 150

RESULT 17
US-09-444-728-2
; Sequence 2, Application US/09444728
; Patent No. 6468799
; GENERAL INFORMATION:
; APPLICANT: Burrell, Michael M.
; TITLE OF INVENTION: Genetically Modified Plants with Altered Starch
; FILE REFERENCE: 9341-019
; CURRENT APPLICATION NUMBER: US/09/444,728
; CURRENT FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Pea ssu TP
; OTHER INFORMATION: linked to E. coli glgC16 CDS
US-09-444-728-2

Query Match 86.2%; Score 25; DB 4; Length 488;
Best Local Similarity 83.3%; Pred. No. 7.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAAAPF 6
11111
Db 22 AAVAPF 27

RESULT 18
US-08-468-700-37
; Sequence 37, Application US/08468700
; Patent No. 5736499
; GENERAL INFORMATION:
; APPLICANT: COLIN MITCHINSON
; APPLICANT: CAROL A. REQUADT
; APPLICANT: TRACI H. ROOP
; APPLICANT: LEIF P. SOLHEIM
; TITLE OF INVENTION: MUTANT ALPHA-AMYLASE
; NUMBER OF SEQUENCES: 40


```

CORRESPONDENCE ADDRESS:
ADDRESSER: Genencor International
STREET: 180 Kimball Way
CITY: South San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,700
FILING DATE: 06-JUN-1995
CLASSIFICATION: 252
ATTORNEY/AGENT INFORMATION:
NAME: Stone, Christopher
REGISTRATION NUMBER: 35,696
REFERENCE/DOCKET NUMBER: GC275
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 742-7555
TELEFAX: (415) 742-7217
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-468-700-37

Query Match      86.2%; Score 25; DB 1; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AAAPF 6
        11111
Db      34 AAAPF 38

RESULT 19
US-08-645-971-5
; Sequence 5, Application US/08645971
; Patent No. 5763385
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Modified Alpha-Amylases Having Altered
; TITLE OF INVENTION: Calcium Binding Properties
; NUMBER OF SEQUENCES: 5
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,971
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Bacillus stearothermophilus
US-08-645-971-5

Query Match      86.2%; Score 25; DB 1; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AAAPF 6
```

```

Db      34 AAAPF 38

RESULT 20
US-08-468-220-35
; Sequence 35, Application US/08468220
; Patent No. 5824532
; GENERAL INFORMATION:
; APPLICANT: Antrim, Richard L.
; APPLICANT: Barnett, Christopher
; APPLICANT: Mitchinson, Colin
; APPLICANT: Power, Scott D.
; APPLICANT: Reguadt, Carol
; APPLICANT: Solheim, Leif P.
; TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genencor International, Inc.
; STREET: 180 Kimball Way
; CITY: South San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 Inch, 1.44 Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,220
; FILING DATE: 06-JUN-95
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/194,664
; FILING DATE: 10-FEB-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/016,395
; FILING DATE: 11-FEB-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Stone, Christopher L.
; REGISTRATION NUMBER: 35,696
; REFERENCE/DOCKET NUMBER: GC220D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 742-7555
; TELEFAX: (415) 742-7217
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-220-35

Query Match      86.2%; Score 25; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AAAPF 6
        11111
Db      34 AAAPF 38

RESULT 21
US-08-468-698-35
; Sequence 35, Application US/08468698
; Patent No. 5849549
; GENERAL INFORMATION:
; APPLICANT: Antrim, Richard L.
; APPLICANT: Barnett, Christopher
; APPLICANT: Mitchinson, Colin
; APPLICANT: Power, Scott D.
```

APPLICANT: Reguadt, Carol
APPLICANT: Solheim, Lelf P.
TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 180 Kimball Way
CITY: South San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,698
FILING DATE: 06-JUN-95
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/194,664
FILING DATE: 10-FEB-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/016,395
FILING DATE: 11-FEB-93
ATTORNEY/AGENT INFORMATION:
NAME: Stone, Christopher L.
REGISTRATION NUMBER: 35,696
REFERENCE/DOCKET NUMBER: GC220D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 742-7555
TELEFAX: (415) 742-7217
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-698-35

Query Match 86.2%; Score 25; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
DB 34 AAAPF 38

RESULT 22
US-08-704-706A-37
Sequence 37, Application US/08704706A
Patent No. 5958739
GENERAL INFORMATION:
APPLICANT: GENENCOR INTERNATIONAL, INC.
APPLICANT: COLIN MITCHINSON
APPLICANT: ANTHONY G. DAY
APPLICANT: ANDREW SHAW
TITLE OF INVENTION: MUTANT ALPHA-AMYLASE
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1013
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/704,706A
FILING DATE: February 20, 1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stone, Christopher
REGISTRATION NUMBER: 35,696
REFERENCE/DOCKET NUMBER: GC275-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 846-7555
TELEFAX: (650) 845-6405
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-704-706A-37

Query Match 86.2%; Score 25; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
DB 34 AAAPF 38

RESULT 23
US-08-890-383-6
Sequence 6, Application US/08890383
Patent No. 6008026
GENERAL INFORMATION:
APPLICANT: Anthony G. Day
TITLE OF INVENTION: MUTANT ALPHA-AMYLASE HAVING INTRODUCED
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1013
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/890,383
FILING DATE: To Be Assigned
ATTORNEY/AGENT INFORMATION:
NAME: Stone, Christopher
REGISTRATION NUMBER: 35,696
REFERENCE/DOCKET NUMBER: GC377
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 846-7555
TELEFAX: (650) 845-6504
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-890-383-6

Query Match 86.2%; Score 25; DB 3; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
11111
Db 34 AAAPF 38

RESULT 24

US-08-914-679A-6
; Sequence 6, Application US/08914679A
; Patent No. 6080568
; GENERAL INFORMATION:
; APPLICANT: Anthony G. Day
; APPLICANT: Barbara A. Swanson
; TITLE OF INVENTION: MUTANT ALPHA-AMYLASE COMPRISING MODIFICATION
; TITLE OF INVENTION: AT RESIDUES CORRESPONDING TO A210, H405 AND/OR
; TITLE OF INVENTION: T412 IN BACILLUS LICHENIFORMIS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genencor International, Inc.
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/914,679A
; FILING DATE: To Be Assigned
; ATTORNEY/AGENT INFORMATION:
; NAME: Stone, Christopher
; REGISTRATION NUMBER: 35,696
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 845-6504
; TELEFAX: (650) 845-6504
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-914-679A-6

Query Match 86.2%; Score 25; DB 3; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
11111
Db 34 AAAPF 38

RESULT 25

US-08-985-659-38
; Sequence 38, Application US/08985659
; Patent No. 6211134
; GENERAL INFORMATION:
; APPLICANT: Caldwell, Robert M
; APPLICANT: ROPP, Traci M
; APPLICANT: Mitchinson, Colin
; TITLE OF INVENTION: Mutant Alpha-Amylase
; FILE REFERENCE: A-69396/DJB/DAV/JJD
; CURRENT APPLICATION NUMBER: US/08/985,659
; CURRENT FILING DATE: 1997-12-09
; PRIOR APPLICATION NUMBER: 08/645,971
; PRIOR FILING DATE: 1996-05-14
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 38
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-08-985-659-38

Query Match 86.2%; Score 25; DB 4; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
11111
Db 34 AAAPF 38

RESULT 26

US-08-194-664A-35
; Sequence 35, Application US/08194664A
; Patent No. 6297037
; GENERAL INFORMATION:
; APPLICANT: Antlim, Richard L.
; APPLICANT: Barnett, Christopher
; APPLICANT: Mitchinson, Colin
; APPLICANT: Power, Scott D.
; APPLICANT: Reguadt, Carol
; TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genencor International, Inc.
; STREET: 180 Kimball Way
; CITY: South San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,664A
; FILING DATE: 10-FEB-94
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stone, Christopher L.
; REGISTRATION NUMBER: 35,696
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 742-7555
; TELEFAX: (415) 742-7217
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-194-664A-35

Query Match 86.2%; Score 25; DB 4; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
11111
Db 34 AAAPF 38

RESULT 27

PCT-US94-01553A-35
; Sequence 35, Application PC/TUS9401553A
; GENERAL INFORMATION:

APPLICANT: GENENCOR INTERNATIONAL, INC.
TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 180 Kimball Way
CITY: South San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01553A
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Horn, Margaret A.
REGISTRATION NUMBER: 33,401
REFERENCE/DOCKET NUMBER: GC220-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 742-7217
TELEFAX: (415) 742-7536
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-01553A-35

Query Match 86.2%; Score 25; DB 5; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
Db 34 AAAPF 38

RESULT 28
PCT-US95-10426-35
Sequence 35, Application PC/TUS9510426
GENERAL INFORMATION:
APPLICANT: GENENCOR INTERNATIONAL, INC.
TITLE OF INVENTION: An Improved Cleaning Composition
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International
STREET: 180 Kimball Way
CITY: South San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10426
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: STONE, Christopher L.
REGISTRATION NUMBER: 33,401
REFERENCE/DOCKET NUMBER: GC220-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 742-7536

TELEFAX: (415) 742-7217
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-10426-35

Query Match 86.2%; Score 25; DB 5; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
Db 34 AAAPF 38

RESULT 29
US-08-720-899-6
Sequence 6, Application US/08720899
Patent No. 5753460
GENERAL INFORMATION:
APPLICANT: Bisgaard-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
APPLICANT: Svendsen, Allan
APPLICANT: Thellersen, Marianne
APPLICANT: Van der Zee, Pia
TITLE OF INVENTION: AMYLASE VARIANTS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,899
FILING DATE: 10-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/343,804
FILING DATE: 22-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4054.214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 549 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-720-899-6

Query Match 86.2%; Score 25; DB 1; Length 549;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
Db 34 AAAPF 38

RESULT 30
US-08-459-610-6
; Sequence 6, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58010430 No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 549 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-610-6

Query Match 86.2%; Score 25; DB 1; Length 549;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
Db 34 AAAPF 38

RESULT 31
US-08-343-804-6
; Sequence 6, Application US/08343804
; Patent No. 5830837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58308370 No. 5830837disk of No. 5830837th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York

COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,804
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 549 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-343-804-6

Query Match 86.2%; Score 25; DB 2; Length 549;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
Db 34 AAAPF 38

RESULT 32
US-08-687-399-6
; Sequence 6, Application US/08687399
; Patent No. 5928381
; GENERAL INFORMATION:
; APPLICANT: Toft, Annette H.
; APPLICANT: Marcher, Dorthie
; APPLICANT: Pedersen, Hanne H.
; APPLICANT: Nilsson, Thomas E.
; TITLE OF INVENTION: A Combined Desizing and Bleaching
; TITLE OF INVENTION: Process
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59283810 No. 5928381disk of No. 5928381th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,399
; FILING DATE:
; CLASSIFICATION: 008
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4127.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 549 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-687-399-6

Query Match 86.2%; Score 25; DB 2; Length 549;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
|||||
Db 34 AAAPF 38

RESULT 33
US-08-600-908A-6
Sequence 6, Application US/08600908A
Patent No. 5989169

GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59891690 No. 5989169disk of No. 5989169th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,908A
FILING DATE: 13-FEB-1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 549 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-600-908A-6

Query Match 86.2%; Score 25; DB 2; Length 549;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
|||||
Db 34 AAAPF 38

RESULT 34
US-08-683-838A-6
Sequence 6, Application US/08683838A
Patent No. 6022724

GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants

NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60227240 No. 6022724disk of No. 6022724th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,838A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/600,908
FILING DATE: 13-FEB-1996

ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 549 amino acid
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-838A-6

Query Match 86.2%; Score 25; DB 3; Length 549;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
|||||
Db 34 AAAPF 38

RESULT 35
US-09-636-252A-6

Sequence 6, Application US/09636252A
Patent No. 6440716

GENERAL INFORMATION:

APPLICANT: Svendsen, Allan

APPLICANT: Bisgard-Frantzen, Henrik

APPLICANT: Borchert, Torben Vedel

TITLE OF INVENTION: Alpha-Amylase Mutants

FILE REFERENCE: 0776/1F216-US2

CURRENT APPLICATION NUMBER: US/09/636,252A

PRIOR FILING DATE: 2000-08-10

PRIOR FILING DATE: 1996-07-18

NUMBER OF SEQ ID NOS: 16

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 6

LENGTH: 549

TYPE: PRT

ORGANISM: B. stearothermophilus
US-09-636-252A-6

Query Match 86.2%; Score 25; DB 4; Length 549;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
|||||
Db 34 AAAPF 38

```
RESULT 36
US-09-313-677-17
; Sequence 17, Application US/09313677
; Patent No. 6300115
; GENERAL INFORMATION:
; APPLICANT: Teague, W. Martin
; APPLICANT: Brumm, Phillip J.
; APPLICANT: Allen, Larry N.
; APPLICANT: Brikun, Igor A.
; TITLE OF INVENTION: Pululanase Expression Constructs Containing Alpha
; TITLE OF INVENTION: Amylase Promoter and Leader Sequences
; FILE REFERENCE: Pululanase Constructs
; CURRENT APPLICATION NUMBER: US/09/313,677
; CURRENT FILING DATE: 1999-05-18
; EARLIER APPLICATION NUMBER: 60/122,065
; EARLIER FILING DATE: 1998-05-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 17
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-313-677-17

Query Match      86.2%; Score 25; DB 4; Length 967;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AAAPF 6
        |||||
DB      34 AAAPF 38

RESULT 37
US-09-171-461-14
; Sequence 14, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schafner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800000
; CURRENT APPLICATION NUMBER: US/09/171,461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 14
; LENGTH: 984
; TYPE: PRT
; ORGANISM: CELO Virus
; FEATURE:
; OTHER INFORMATION: Position: 23680..26634 /gene: L4 /product: L4 100K
US-09-171-461-14

Query Match      86.2%; Score 25; DB 4; Length 984;
Best Local Similarity 83.3%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAPF 6
        |||||
DB      644 AAAPF 649

RESULT 38
US-08-459-568-2
; Sequence 2, Application US/08459568
```

```
; Patent No. 5811304
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESS: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,568
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,411
; FILING DATE: 06-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1706 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-459-568-2

Query Match      86.2%; Score 25; DB 2; Length 1706;
Best Local Similarity 83.3%; Pred. No. 2.7e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAPF 6
        |||||
DB      1694 AAATPF 1699

RESULT 39
US-08-399-411-2
; Sequence 2, Application US/08399411
; Patent No. 5831008
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESS: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/399,411
; FILING DATE: 06-MAR-1995
```

```

; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1706 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-399-411-2

```

```

Query Match      86.2%; Score 25; DB 2; Length 1706;
Best Local Similarity 83.3%; Pred. No. 2.7e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 AAAPF 6
    ||| ||
Db 1694 AAATPF 1699

```

```

RESULT 40
US-08-516-859A-2
; Sequence 2, Application US/08516859A
; Patent No. 6069231
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/516,859A
; FILING DATE: 18-AUG-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,411
; FILING DATE: 06-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/292,683
; FILING DATE: 18-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1776
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1706 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-516-859A-2

```

```

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAPF 6
    ||| ||
Db 1694 AAATPF 1699

```

Search completed: December 6, 2002, 13:31:54
Job time : 17 secs

```

Query Match      86.2%; Score 25; DB 3; Length 1706;
Best Local Similarity 83.3%; Pred. No. 2.7e+03;

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:30:14 ; Search time 7.33333 Seconds
(without alignments)
13.289 Million cell updates/sec

Title: US-10-033-526-4
Perfected score: 29
Sequence: 1 AAAAPF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries.

Database: Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------------------|-------------------|
| 1 | 29 | 100.0 | 6 | 12 US-10-033-526-4 | Sequence 4, Appli |
| 2 | 29 | 100.0 | 405 | 10 US-09-740-288A-18 | Sequence 18, Appl |
| 3 | 29 | 100.0 | 410 | 9 US-10-078-650-14 | Sequence 14, Appl |
| 4 | 29 | 100.0 | 482 | 9 US-10-078-650-2 | Sequence 2, Appli |
| 5 | 29 | 100.0 | 484 | 9 US-10-078-650-12 | Sequence 12, Appl |
| 6 | 26 | 89.7 | 9 | 1 US-08-821-739A-13 | Sequence 13, Appl |
| 7 | 26 | 89.7 | 9 | 1 US-08-821-739A-25 | Sequence 25, Appl |
| 8 | 26 | 89.7 | 597 | 10 US-09-793-306-146 | Sequence 146, App |
| 9 | 26 | 89.7 | 4613 | 9 US-09-860-846-31 | Sequence 31, Appl |
| 10 | 26 | 89.7 | 4613 | 9 US-09-861-289-31 | Sequence 31, Appl |
| 11 | 26 | 89.7 | 11877 | 10 US-09-860-846-6 | Sequence 6, Appli |
| 12 | 26 | 89.7 | 11877 | 10 US-09-861-289-6 | Sequence 566, App |
| 13 | 25 | 86.2 | 51 | 10 US-09-925-297-566 | Sequence 1206, Ap |
| 14 | 25 | 86.2 | 89 | 10 US-09-925-301-1206 | Sequence 1090, Ap |
| 15 | 25 | 86.2 | 129 | 10 US-09-764-864-1090 | Sequence 1511, Ap |
| 16 | 25 | 86.2 | 136 | 10 US-09-764-864-1511 | Sequence 824, App |
| 17 | 25 | 86.2 | 139 | 10 US-09-764-847-824 | Sequence 154, App |
| 18 | 25 | 86.2 | 247 | 10 US-09-923-779-154 | Sequence 411, App |
| 19 | 25 | 86.2 | 254 | 10 US-09-741-669-411 | |

| | | | | | |
|----|----|------|------|------------------------|-------------------|
| 20 | 25 | 86.2 | 307 | 12 US-10-062-254-244 | Sequence 244, App |
| 21 | 25 | 86.2 | 436 | 12 US-10-062-254-248 | Sequence 248, App |
| 22 | 25 | 86.2 | 456 | 9 US-09-989-920-175 | Sequence 175, App |
| 23 | 25 | 86.2 | 456 | 10 US-09-993-811-2 | Sequence 2, Appli |
| 24 | 25 | 86.2 | 456 | 10 US-09-974-712-2 | Sequence 4, Appli |
| 25 | 25 | 86.2 | 476 | 9 US-10-058-820-4 | Sequence 3, Appli |
| 26 | 25 | 86.2 | 553 | 9 US-10-058-820-3 | Sequence 37, Appl |
| 27 | 25 | 86.2 | 580 | 9 US-09-928-457-37 | Sequence 401, App |
| 28 | 25 | 86.2 | 770 | 10 US-09-741-669-401 | Sequence 14, Appl |
| 29 | 25 | 86.2 | 984 | 10 US-09-970-711-14 | Sequence 2, Appli |
| 30 | 25 | 86.2 | 1068 | 10 US-09-858-081-2 | Sequence 2, Appli |
| 31 | 25 | 86.2 | 1133 | 10 US-09-858-068-2 | Sequence 44, Appl |
| 32 | 25 | 86.2 | 93 | 10 US-09-908-322-44 | Sequence 508, App |
| 33 | 25 | 86.2 | 144 | 10 US-09-867-550-508 | Sequence 668, App |
| 34 | 24 | 82.8 | 168 | 10 US-09-864-761-49050 | Sequence 620, App |
| 35 | 24 | 82.8 | 190 | 10 US-09-867-550-668 | Sequence 12716, A |
| 36 | 24 | 82.8 | 248 | 10 US-09-764-853-620 | Sequence 9, Appli |
| 37 | 24 | 82.8 | 290 | 10 US-09-815-242-12716 | Sequence 9, Appli |
| 38 | 24 | 82.8 | 315 | 9 US-10-041-006A-9 | Sequence 9, Appli |
| 39 | 24 | 82.8 | 315 | 12 US-10-040-655-9 | Sequence 9, Appli |
| 40 | 24 | 82.8 | 316 | 12 US-10-041-400A-9 | Sequence 9, Appli |
| 41 | 24 | 82.8 | 316 | 12 US-10-041-264A-9 | Sequence 9, Appli |
| 42 | 24 | 82.8 | 316 | 12 US-10-042-091A-9 | Sequence 8, Appli |
| 43 | 24 | 82.8 | 389 | 9 US-10-040-803-8 | Sequence 2, Appli |
| 44 | 24 | 82.8 | 329 | 10 US-09-767-770A-2 | Sequence 5076, Ap |
| 45 | 24 | 82.8 | 400 | 10 US-09-815-242-5076 | |

ALIGNMENTS

RESULT 1
US-10-033-526-4
; Sequence 4, Application US/10033526
; Patent No. US20020147999A1
; GENERAL INFORMATION:
; APPLICANT: Robert W. Mahley
; APPLICANT: Yadong Huang
; TITLE OF INVENTION: Methods of Treating Disorders Related to
; FILE REFERENCE: UCA1217
; CURRENT APPLICATION NUMBER: US/10/033,526
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 60/245,737
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-033-526-4

Query Match 100.0%; Score 29; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 1 AAAAPF 6

RESULT 2
US-09-740-288A-18
; Sequence 18, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil

```
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: B81429 US NA
; CURRENT APPLICATION NUMBER: US/09/740,288A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Hordeum vulgare
; US-09-740-288A-18
```

```
Query Match          100.0%; Score 29; DB 10; Length 405;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAAAPF 6
        |||||
Db       11 AAAAPF 16
```

```
RESULT 3
US-10-078-650-14
; Sequence 14, Application US/10078650
; Patent No. US20020169301A1
; GENERAL INFORMATION:
; APPLICANT: Fujimoto, Katsumi
; APPLICANT: Shin, Mei
; APPLICANT: Kato, Yukio
; TITLE OF INVENTION: NOVEL bHLH TYPE TRANSCRIPTION FACTOR
; FILE REFERENCE: 06501-101001
; CURRENT APPLICATION NUMBER: US/10/078,650
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/JP00/03991
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: JP 11-233286
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-078-650-14
```

```
Query Match          100.0%; Score 29; DB 9; Length 410;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAAAPF 6
        |||||
Db       321 AAAAPF 326
```

```
RESULT 4
US-10-078-650-2
; Sequence 2, Application US/10078650
; Patent No. US20020169301A1
; GENERAL INFORMATION:
; APPLICANT: Fujimoto, Katsumi
; APPLICANT: Shin, Mei
; APPLICANT: Kato, Yukio
; TITLE OF INVENTION: NOVEL bHLH TYPE TRANSCRIPTION FACTOR
; FILE REFERENCE: 06501-101001
; CURRENT APPLICATION NUMBER: US/10/078,650
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/JP00/03991
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: JP 11-233286
```

```
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-078-650-2
```

```
Query Match          100.0%; Score 29; DB 9; Length 482;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAAAPF 6
        |||||
Db       380 AAAAPF 385
```

```
RESULT 5
US-10-078-650-12
; Sequence 12, Application US/10078650
; Patent No. US20020169301A1
; GENERAL INFORMATION:
; APPLICANT: Fujimoto, Katsumi
; APPLICANT: Shin, Mei
; APPLICANT: Kato, Yukio
; TITLE OF INVENTION: NOVEL bHLH TYPE TRANSCRIPTION FACTOR
; FILE REFERENCE: 06501-101001
; CURRENT APPLICATION NUMBER: US/10/078,650
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/JP00/03991
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: JP 11-233286
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-078-650-12
```

```
Query Match          100.0%; Score 29; DB 9; Length 484;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAAAPF 6
        |||||
Db       382 AAAAPF 387
```

```
RESULT 6
US-08-821-739A-13
; Sequence 13, Application US/08821739A
; Patent No. US20020168374A1
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding Peptides and Their Uses
; FILE REFERENCE: 2060.005000A
; CURRENT APPLICATION NUMBER: US/08/821,739A
; CURRENT FILING DATE: 1999-03-20
; PRIOR APPLICATION NUMBER: 60/013,833
; PRIOR FILING DATE: 1996-03-21
; PRIOR APPLICATION NUMBER: 08/589,107
; PRIOR FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 08/451,913
; PRIOR FILING DATE: 1995-05-26
; PRIOR APPLICATION NUMBER: 08/347,610
; PRIOR FILING DATE: 1994-12-01
```


; PRIOR APPLICATION NUMBER: 08/186,266
; PRIOR FILING DATE: 1994-01-25
; PRIOR APPLICATION NUMBER: 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: 08/027,746
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: 07/926,666
; PRIOR FILING DATE: 1992-08-07
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 13
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-08-821-739A-13

Query Match 89.7%; Score 26; DB 1; Length 9;
Best Local Similarity 83.3%; Pred. No. 8.6e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|1111:
Db 4 AAAAPY 9

RESULT 7
US-08-821-739A-25
; Sequence 25, Application US/08821739A
; Patent No. US20020168374A1
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding Peptides and Their Uses
; FILE REFERENCE: 2060.005000A
; CURRENT APPLICATION NUMBER: US/08/821,739A
; CURRENT FILING DATE: 1999-03-20
; PRIOR APPLICATION NUMBER: 60/013,833
; PRIOR FILING DATE: 1996-03-21
; PRIOR APPLICATION NUMBER: 08/589,107
; PRIOR FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 08/451,913
; PRIOR FILING DATE: 1995-05-26
; PRIOR APPLICATION NUMBER: 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: 08/186,266
; PRIOR FILING DATE: 1994-01-25
; PRIOR APPLICATION NUMBER: 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: 08/027,746
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: 07/926,666
; PRIOR FILING DATE: 1992-08-07
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 25
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-08-821-739A-25

Query Match 89.7%; Score 26; DB 1; Length 9;
Best Local Similarity 83.3%; Pred. No. 8.6e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|1111:
Db 4 AAAAPY 9

RESULT 8
US-09-793-306-146
; Sequence 146, Application US/097933306
; Patent No. US20020098200A1
; GENERAL INFORMATION:
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Skeiky, Yasir
; APPLICANT: Owendale, Pamela
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compounds and Methods for Diagnosis and Immunotherapy
; FILE REFERENCE: 014058-008740US
; CURRENT APPLICATION NUMBER: US/09/793,306
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: US 60/185,037
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 60/223,828
; PRIOR FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 164
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 146
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mttc#3-His
; US-09-793-306-146

Query Match 89.7%; Score 26; DB 10; Length 597;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|1111:
Db 74 AAAAPY 79

RESULT 9
US-09-860-846-31
; Sequence 31, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-860-846-31

Query Match 89.7%; Score 26; DB 9; Length 4613;
Best Local Similarity 83.3%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|1111:
Db 3787 AAAAPY 3792

```
RESULT 10
US-09-861-289-31
; Sequence 31, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-861-289-31
```

```
Query Match      89.7%; Score 26; DB 10; Length 4613;
Best Local Similarity 83.3%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAAAPF 6
        |||||:
Db      4123 AAAAPY 3792
```

```
RESULT 11
US-09-860-846-6
; Sequence 6, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 11877
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-6
```

```
Query Match      89.7%; Score 26; DB 9; Length 11877;
Best Local Similarity 83.3%; Pred. No. 5.5e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAAAPF 6
        |||||:
Db      4123 AAAAPY 4128
```

```
RESULT 12
US-09-861-289-6
; Sequence 6, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
```

```
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 11877
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-861-289-6
```

```
Query Match      89.7%; Score 26; DB 10; Length 11877;
Best Local Similarity 83.3%; Pred. No. 5.5e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AAAAPF 6
        |||||:
Db      4123 AAAAPY 4128
```

```
RESULT 13
US-09-925-297-566
; Sequence 566, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 566
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (35)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (36)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (48)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-566
```

```
Query Match      86.2%; Score 25; DB 10; Length 51;
Best Local Similarity 83.3%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 AAAAPF 6
        |||||:
Db      19 AVAAPF 24
```

```
RESULT 14
US-09-925-301-1206
; Sequence 1206, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
```

```
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1206
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1206

Query Match      86.2%; Score 25; DB 10; Length 89;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AAAPF 6
        |||||
Db      77 AAAPF 81

RESULT 15
US-09-764-864-1090
; Sequence 1090, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1090
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (27)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (80)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1090

Query Match      86.2%; Score 25; DB 10; Length 129;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAPF 6
        |||||
Db      6 AACAPF 11

RESULT 16
US-09-764-864-1511
; Sequence 1511, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17

Prior application data removed - consult PALM or file wrapper

; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1511
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (27)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (80)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1511

Query Match      86.2%; Score 25; DB 10; Length 129;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAPF 6
        |||||
Db      6 AACAPF 11

RESULT 17
US-09-764-847-824
; Sequence 824, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 824
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (106)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (135)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-824

Query Match      86.2%; Score 25; DB 10; Length 136;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAPF 6
        |||||
Db      26 AAXAPF 31
```

RESULT 18
US-09-923-779-154
; Sequence 154, Application US/09923779
; Patent No. US20020076721A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.553
; CURRENT APPLICATION NUMBER: US/09/923,779
; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 154
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-779-154

Query Match 86.2%; Score 25; DB 10; Length 247;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAPF 6
1 ||||
Db 13 AVAAPF 18

RESULT 19
US-09-741-669-411
; Sequence 411, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; TITLE OF INVENTION: proliferation of E. coli
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 411
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-411

Query Match 86.2%; Score 25; DB 10; Length 254;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAPF 6
1 ||||
Db 60 AVAAPF 65

RESULT 20
US-10-062-254-244
; Sequence 244, Application US/10062254
; Patent No. US20020138882A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Fang, Yiwen

; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jian-Ming
; APPLICANT: Li, Zhongsen
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Niu, Xiping
; APPLICANT: Odell, Joan
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; APPLICANT: Zheng, Peizhong
; APPLICANT: Zhu, Qun
; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved in Plant Metabolism
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/062,254
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/630,346
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/146511
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/156006
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/156899
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/157287
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/169767
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/171054
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/172958
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/171515
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 60/173535
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 244
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Zea mays
US-10-062-254-244

Query Match 86.2%; Score 25; DB 12; Length 307;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
1 ||||
Db 279 AAAPF 283

RESULT 21
US-10-062-254-248
; Sequence 248, Application US/10062254
; Patent No. US20020138882A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Fang, Yiwen
; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jian-Ming
; APPLICANT: Li, Zhongsen
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Niu, Xiping
; APPLICANT: Odell, Joan
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; APPLICANT: Zheng, Peizhong
; APPLICANT: Zhu, Qun
; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved in Plant Metabolism

```

; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/062,254
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/630,346
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/146511
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/156006
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/156899
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/157287
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/169767
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/171054
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/172958
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/171515
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 60/173535
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 248
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Zea mays
US-10-062-254-248

Query Match      86.2%; Score 25; DB 12; Length 436;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AAAAPF 6
      |||||
Db      408 AAAAPF 412

RESULT 22
US-09-989-920-175
; Sequence 175, Application US/09989920
; Patent No. US20020172957A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 175
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-175

Query Match      86.2%; Score 25; DB 9; Length 456;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAAPF 6
      |||||
Db      181 AAAAPF 186

; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/062,254
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/630,346
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/146511
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/156006
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/156899
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/157287
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/169767
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/171054
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/172958
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/171515
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 60/173535
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 248
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Zea mays
US-10-062-254-248

Query Match      86.2%; Score 25; DB 12; Length 436;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AAAAPF 6
      |||||
Db      408 AAAAPF 412

RESULT 23
US-09-993-811-2
; Sequence 2, Application US/09993811
; Patent No. US20020119476A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Nucleotide Sequences
; FILE REFERENCE: 00248
; CURRENT APPLICATION NUMBER: US/09/993,811
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-993-811-2

Query Match      86.2%; Score 25; DB 10; Length 456;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAAPF 6
      |||||
Db      181 AAAAPF 186

RESULT 24
US-09-974-712-2
; Sequence 2, Application US/09974712
; Patent No. US20020119540A1
; GENERAL INFORMATION:
; APPLICANT: Fiddler, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020119540A1 Human Ion Channel Protein and Polynucleo
; FILE REFERENCE: LEX-0251-USA
; CURRENT APPLICATION NUMBER: US/09/974,712
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,623
; PRIOR FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 456
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-974-712-2

Query Match      86.2%; Score 25; DB 10; Length 456;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAAPF 6
      |||||
Db      181 AAAAPF 186

RESULT 25
US-10-058-820-4
; Sequence 4, Application US/10058820
; Patent No. US20020155479A1
; GENERAL INFORMATION:
; APPLICANT: Bogan, Jonathan S.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: Expression Cloning Method
; FILE REFERENCE: 0399.2025-002
; CURRENT APPLICATION NUMBER: US/10/058,820
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/325,651
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/298,963
```



```
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/264,816
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-058-820-4
```

```
Query Match      86.2%; Score 25; DB 9; Length 476;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 AAAPF 6
      11111
Db      161 AAAPF 165
```

RESULT 26

```
US-10-058-820-3
; Sequence 3, Application US/10058820
; Patent No. US20020155479A1
; GENERAL INFORMATION:
; APPLICANT: Bogan, Jonathan S.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: Expression Cloning Method
; FILE REFERENCE: 0399.2025-002
; CURRENT APPLICATION NUMBER: US/10/058,820
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/325,651
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/298,963
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/264,816
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-058-820-3
```

```
Query Match      86.2%; Score 25; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 AAAPF 6
      11111
Db      238 AAAPF 242
```

RESULT 27

```
US-09-928-457-37
; Sequence 37, Application US/09928457
; Patent No. US20020164603A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA, specific proteins and peptides
; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method
; TITLE OF INVENTION: for obtaining them and their biological application.
; NUMBER OF SEQUENCES: 99
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (OEB)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/928,457
; FILING DATE: 2001-08-14
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 09/214,759
; FILING DATE: 199-12-10
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 580 acids amin,s
; TYPE: acide amin,
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..580
US-09-928-457-37
```

```
Query Match      86.2%; Score 25; DB 9; Length 580;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 AAAPF 6
      11111
Db      439 AAAPF 443
```

RESULT 28

```
US-09-741-669-401
; Sequence 401, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; TITLE OF INVENTION: proliferation of E. coli
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 401
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-401
```

```
Query Match      86.2%; Score 25; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 AAAPF 6
      11111
Db      15 AAAPF 19
```

RESULT 29

```
US-09-970-711-14
; Sequence 14, Application US/09970711
; Patent No. US20020081279A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbauner, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800001
; CURRENT APPLICATION NUMBER: US/09/970,711
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/171,461
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: PCT/EP97/01944
; PRIOR FILING DATE: 1997-04-18
```

```
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 14
; LENGTH: 984
; TYPE: PRT
; ORGANISM: CELO Virus
; FEATURE:
; OTHER INFORMATION: Position: 23680..26634 /gene: L4 /product: L4 100K
US-09-970-711-14

Query Match      86.2%; Score 25; DB 10; Length 984;
Best Local Similarity 83.3%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAAPF 6
Db      644 AVAAPF 649

RESULT 30
US-09-858-081-2
; Sequence 2, Application US/09858081
; Patent No. US20020072490A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Liebermann, Rosana
; APPLICANT: Cook, William James
; TITLE OF INVENTION: 33428, A NOVEL HUMAN METALLOPROTEASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-049001
; CURRENT APPLICATION NUMBER: US/09/858, 081
; CURRENT FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/204,159
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1068
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-858-081-2

Query Match      86.2%; Score 25; DB 10; Length 1068;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AAAPF 6
Db      818 AAAPF 822

RESULT 31
US-09-858-068-2
; Sequence 2, Application US/09858068
; Patent No. US20020076778A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Liebermann, Rosana
; APPLICANT: Cook, William James
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: 33428, A NOVEL HUMAN METALLOPROTEASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-057001
; CURRENT APPLICATION NUMBER: US/09/858,068
; CURRENT FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/204,159
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: 60/204,160
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 2

; LENGTH: 1133
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-858-068-2

Query Match      86.2%; Score 25; DB 10; Length 1133;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AAAPF 6
Db      883 AAAPF 887

RESULT 32
US-09-908-322-44
; Sequence 44, Application US/09908322
; Patent No. US20020107194A1
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos Manuel Pinto
; APPLICANT: Lewis, Julian Hart
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Gray, Grace
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF
; TITLE OF INVENTION: VERTEBRATE DELTA GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennile & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/908,322
; FILING DATE: 17-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/981,392
; FILING DATE: 22-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-908-322-44

Query Match      82.8%; Score 24; DB 10; Length 93;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAAAPF 6
Db      22 AAAAPW 27
```

```
RESULT 33
US-09-867-550-508
; Sequence 508, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 508
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (137)
; OTHER INFORMATION: wherein Xaa may be any one of Leu or Phe
US-09-867-550-508

Query Match      82.8%; Score 24; DB 10; Length 144;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAAPF 6
      ||| ||
Db      51 AAAAPF 56

RESULT 34
US-09-864-761-49050
; Sequence 49050, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 49050
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO 282214.22
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.6
; OTHER INFORMATION: EST_HUMAN HIT: A1146468.1, EVALUE 4.00e-03
US-09-864-761-49050

Query Match      82.8%; Score 24; DB 10; Length 168;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAAPF 6
      ||| |||
Db      35 AAPAPF 40

RESULT 35
US-09-867-550-668
; Sequence 668, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 668
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-09-867-550-668

Query Match      82.8%; Score 24; DB 10; Length 190;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 AAAAPF 6
      ||| |||
Db      87 AAAKPF 92

RESULT 36
US-09-764-853-620
; Sequence 620, Application US/09764853
```

```
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206
; CURRENT APPLICATION NUMBER: US/09/764,853
; PRIOR application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 620
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-853-620
```

```
Query Match      82.8%; Score 24; DB 10; Length 248;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 AAAAPF 6
    | | | | |
Db 60 ARAAPF 65
```

```
RESULT 37
US-09-815-242-12716
; Sequence 12716, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12716
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12716
```

```
Query Match      82.8%; Score 24; DB 10; Length 290;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 AAAAPF 6
    | | | | |
Db 5 AAAAPF 10
```

```
RESULT 38
US-10-041-006A-9
; Sequence 9, Application US/10041006A
; Patent No. US20020168754A1
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/10/041,006A
; CURRENT FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
; OTHER INFORMATION: of Protease T in a zymogen activation construct
US-10-041-006A-9
```

```
Query Match      82.8%; Score 24; DB 9; Length 315;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 AAAAPF 6
    | | | | |
Db 41 ALAAPF 46
```

```
RESULT 39
US-10-040-655-9
; Sequence 9, Application US/10040655
; Patent No. US20020146805A1
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/10/040,655
; CURRENT FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
; OTHER INFORMATION: of Protease T in a zymogen activation construct
US-10-040-655-9
```

```
Query Match      82.8%; Score 24; DB 12; Length 315;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 AAAAPF 6
    | | | | |
Db 41 ALAAPF 46
```

```
RESULT 40
US-10-041-400A-9
; Sequence 9, Application US/10041400A
; Patent No. US20020110895A1
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
```

APPLICANT: Andrade-Gordon, Patricia
TITLE OF INVENTION: DNA Encoding the Human Serine
TITLE OF INVENTION: Protease EOS
FILE REFERENCE: ORT-1031
CURRENT APPLICATION NUMBER: US/10/041,400A
CURRENT FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US/09/387,375
PRIOR FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 316
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Amino acid
OTHER INFORMATION: sequence of EOS zymogen fusion gene
US-10-041-400A-9

Query Match 82.8%; Score 24; DB 12; Length 316;
Best Local Similarity 83.3%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAAPF 6
DB 41 ALAAPF 46

Search completed: December 6, 2002, 13:42:09
Job time : 10.3333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:25:59 ; Search time 15.3333 Seconds
(Without alignments)
37.618 Million cell updates/sec

Title: US-10-033-526-4
Perfect score: 29
Sequence: 1 AAAAPF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pirl:*
2: pirl:*
3: pirl:*
4: pirl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1 | 29 | 100.0 | 102 | 2 | H70898 | probable PE protei |
| 2 | 29 | 100.0 | 130 | 2 | T29498 | hypothetical prote |
| 3 | 29 | 100.0 | 389 | 2 | E95422 | probable transmemb |
| 4 | 29 | 100.0 | 410 | 2 | JC7584 | basic helix-loop-h |
| 5 | 29 | 100.0 | 446 | 2 | A42029 | transcription fact |
| 6 | 29 | 100.0 | 482 | 2 | JC7583 | basic helix-loop-h |
| 7 | 29 | 100.0 | 554 | 2 | A75296 | probable erythromy |
| 8 | 29 | 100.0 | 797 | 2 | A71267 | hypothetical prote |
| 9 | 29 | 100.0 | 1044 | 2 | S01966 | GRPase-activating |
| 10 | 26 | 89.7 | 134 | 2 | A72489 | hypothetical prote |
| 11 | 26 | 89.7 | 161 | 2 | AF2262 | allophycocyanin B |
| 12 | 26 | 89.7 | 192 | 2 | T13106 | minor tail protein |
| 13 | 26 | 89.7 | 206 | 2 | H83587 | conserved hypotet |
| 14 | 26 | 89.7 | 206 | 2 | T36643 | probable integral |
| 15 | 26 | 89.7 | 301 | 2 | C87296 | beta-glucanase lim |
| 16 | 26 | 89.7 | 317 | 2 | C87402 | conserved hypotet |
| 17 | 26 | 89.7 | 329 | 2 | T00873 | hypothetical prote |
| 18 | 26 | 89.7 | 333 | 2 | D70792 | hypothetical prote |
| 19 | 26 | 89.7 | 334 | 2 | T35939 | probable transport |
| 20 | 26 | 89.7 | 400 | 2 | T29121 | hypothetical prote |
| 21 | 26 | 89.7 | 403 | 2 | H70931 | hypothetical prote |
| 22 | 26 | 89.7 | 407 | 2 | S22586 | probable PPE prote |
| 23 | 26 | 89.7 | 409 | 2 | A70647 | homeotic protein E |
| 24 | 26 | 89.7 | 414 | 2 | H90815 | probable PPE prote |
| 25 | 26 | 89.7 | 414 | 2 | D85675 | probable kinase li |
| 26 | 26 | 89.7 | 414 | 2 | AD0645 | ABC transporter in |
| 27 | 26 | 89.7 | 414 | 2 | C64856 | Probable permease |
| 28 | 26 | 89.7 | 416 | 2 | S12541 | evxl protein - mur |
| 29 | 26 | 89.7 | 418 | 2 | F84379 | methylaspartate am |

| | | | | | | |
|----|----|------|-----|---|--------|--------------------|
| 30 | 26 | 89.7 | 434 | 2 | B54843 | nemo, form II - fr |
| 31 | 26 | 89.7 | 434 | 2 | E84879 | probable heme A fa |
| 32 | 26 | 89.7 | 443 | 2 | C70780 | probable PPE prote |
| 33 | 26 | 89.7 | 460 | 2 | T48137 | copper amine oxida |
| 34 | 26 | 89.7 | 463 | 2 | C70931 | probable PPE prote |
| 35 | 26 | 89.7 | 463 | 2 | T13425 | regulatory protein |
| 36 | 26 | 89.7 | 468 | 2 | B70932 | probable PPE prote |
| 37 | 26 | 89.7 | 469 | 2 | T01579 | heme A farnesyltra |
| 38 | 26 | 89.7 | 471 | 1 | S08325 | flavonol 3-O-gluc |
| 39 | 26 | 89.7 | 477 | 2 | A54843 | nemo, form I - fru |
| 40 | 26 | 89.7 | 499 | 2 | AH2761 | penicillin-binding |
| 41 | 26 | 89.7 | 506 | 2 | G97542 | penicillin-binding |
| 42 | 26 | 89.7 | 512 | 2 | E83060 | iron (III)-transpo |
| 43 | 26 | 89.7 | 515 | 1 | DNET02 | NADH2 dehydrogenas |
| 44 | 26 | 89.7 | 521 | 2 | T11710 | probable amino aci |
| 45 | 26 | 89.7 | 552 | 2 | D70604 | probable PPE prote |

ALIGNMENTS

RESULT 1
H70898
Probable PE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: H70898
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A: Reference number: A70500; MUID: 98295987; PMID: 9634230
A: Accession: H70898
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-102 <COL>
A: Cross-references: GB:Z80108; GB:AL123456; NID:g3256012; PIDN:CAB02191.1; PID:e26555
A: Experimental source: strain H37RV
C: Genetics:
A: Gene: PE

Query Match 100.0%; Score 29; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
DB 29 AAAAPF 34

RESULT 2
T29498
hypothetical protein K06B9.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29498
R: Miller, N.; Bradshaw, H.
Submitted to the EMBL Data Library, February 1996
A: Description: The sequence of C. elegans cosmid K06B9.
A: Reference number: Z20628
A: Accession: T29498
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-130 <MIL>
A: Cross-references: EMBL:U50072; PIDN:AAA93448.1; CESP:K06B9.3
C: Genetics:
A: Gene: CESP:K06B9.3
A: Introns: 32/3; 64/1

Query Match 100.0%; Score 29; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 35;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|11111|
Db 2 AAAAPF 7

RESULT 3

E95422

Probable transmembrane transport protein SMA2377 [imported] - Sinorhizobium meliloti (st
C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: E95422

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: E95422

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-389 <KUR>

A:Cross-references: GB:AE006469; PIDN:AAK65943.1; PID:g14524459; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SMA2377

A:Genome: plasmid

Query Match

Best Local Similarity 100.0%; Score 29; DB 2; Length 389;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|11111|

Db 354 AAAAPF 359

RESULT 4

JC7584

basic helix-loop-helix protein, DEC2 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001

C:Accession: JC7584

R:Fujimoto, K.; Shen, M.; Noshiro, M.; Matsubara, K.; Shingu, S.; Honda, K.; Yoshida, E.
Biochem. Biophys. Res. Commun. 280, 164-171, 2001

A:Title: Molecular cloning and characterization of DEC2, a new member of basic helix-loo
A:Reference number: JC7583; MUID:21092582; PMID:11162494

A:Accession: JC7584

A:Molecule type: mRNA

A:Residues: 1-410 <FUJ>

A:Cross-references: DDBJ:AB044090

C:Comment: This protein, a novel member of the DEC subfamily of basic helix-loop-helix p

C:Genetics:

A:Gene: dec2

A:Map position: 6 G2-G3

C:Keywords: transcription factor

F:1-173/Region: highly conserved #status predicted

F:130-173/Domain: Orange #status predicted <ORA>

F:277-344/Region: alanine and glycine-rich #status predicted

Query Match

Best Local Similarity 100.0%; Score 29; DB 2; Length 410;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|11111|
Db 321 AAAAPF 326

RESULT 5

A42029

transcription factor E3 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C:Accession: A42029

R:Roman, C.; Matera, A.G.; Cooper, C.; Artandi, S.; Blain, S.; Ward, D.C.; Calame, K.
Mol. Cell. Biol. 12, 817-827, 1992

A:Title: mTFE3, an X-linked transcriptional activator containing basic helix-loop-hel
A:Reference number: A42029; MUID:92123207; PMID:1732746

A:Accession: A42029

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-446 <ROM>

A:Cross-references: GB:S76673; NID:g243439; PIDN:AAB21130.1; PID:g243440
A:Note: sequence extracted from NCBI backbone (NCBIN:76673, NCBIP:76674)

Query Match

Best Local Similarity 100.0%; Score 29; DB 2; Length 446;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|11111|

Db 10 AAAAPF 15

RESULT 6

JC7583

basic helix-loop-helix protein, DEC2 - human

C:Species: Homo sapiens (man)

C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001

C:Accession: JC7583

R:Fujimoto, K.; Shen, M.; Noshiro, M.; Matsubara, K.; Shingu, S.; Honda, K.; Yoshida,
Biochem. Biophys. Res. Commun. 280, 164-171, 2001

A:Title: Molecular cloning and characterization of DEC2, a new member of basic helix-
A:Reference number: JC7583; MUID:21092582; PMID:11162494

A:Accession: JC7583

A:Molecule type: mRNA

A:Residues: 1-482 <FUJ>

A:Cross-references: DDBJ:AB044088

C:Comment: This protein, a novel member of the DEC subfamily of basic helix-loop-hel

C:Genetics:

A:Gene: dec2

A:Map position: 12p11.23-p12.1.

C:Keywords: transcription factor

F:1-173/Region: highly conserved #status predicted

F:130-173/Domain: Orange #status predicted <ORA>

F:286-411/Region: alanine and glycine-rich #status predicted

Query Match

Best Local Similarity 100.0%; Score 29; DB 2; Length 482;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|11111|

Db 380 AAAAPF 385

RESULT 7

A75296

probable erythromycin esterase - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: A75296

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
.; Shen, M.; Vamathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

```
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: A75296
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-554 <WHI>
A:Cross-references: GB:AE002058; GB:AE000513; NID:g6460059; PIDN:AAF11803.1; PID:g646006
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2257
A:Map position: 1

Query Match
Best Local Similarity 100.0%; Score 29; DB 2; Length 554;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
   |||||
Db 324 AAAAPF 329

RESULT 8
A71267
hypothetical protein TP0900 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 24-Nov-1999
C:Accession: A71267
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwiltson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDcthey, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: A71267
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-797 <COL>
A:Cross-references: GB:AE001259; GB:AE000520; NID:g3323209; PIDN:AAC65868.1; PID:g332322
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0900
C:Superfamily: syphilis spirochete hypothetical protein TP0900

Query Match
Best Local Similarity 100.0%; Score 29; DB 2; Length 797;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
   |||||
Db 512 AAAAPF 517

RESULT 9
S01966
GTPase-activating protein - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 05-Nov-1999
C:Accession: S01966
R:Vogel, U.S.; Dixon, R.A.F.; Schaber, M.D.; Diehl, R.E.; Marshall, M.S.; Scolnick, E.M.
Nature 335, 90-93, 1988
A:Title: Cloning of bovine GAP and its interaction with oncogenic ras p21.
A:Reference number: S01966; MUID:88318957; PMID:2842690
A:Accession: S01966
A:Molecule type: mRNA
A:Residues: 1-1044 <VOG>
A:Cross-references: EMBL:X12602; NID:g384; PIDN:CAA31122.1; PID:g385
A>Note: part of this sequence was confirmed by protein sequencing
C:Superfamily: pleckstrin repeat homology; ras-specific GAP catalytic domain homology; S
F:178-273/Domain: SH2 homology <SH21>
F:283-333/Domain: SH2 homology <SH3>
F:348-442/Domain: SH2 homology <SH22>
F:470-572/Domain: pleckstrin repeat homology <PLK>
F:745-959/Domain: ras-specific GAP catalytic domain homology <GAP>
```

```
Query Match
Best Local Similarity 100.0%; Score 29; DB 2; Length 1044;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
   |||||
Db 45 AAAAPF 50

RESULT 10
A72489
hypothetical protein APE2552 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C:Accession: A72489
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: A72489
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-134 <KAW>
A:Cross-references: DDBJ:AP000064; NID:g5105945; PIDN:BAA81569.1; PID:d1045355; PID:g
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2552
C:Superfamily: Aeropyrum pernix hypothetical protein APE2552

Query Match
Best Local Similarity 89.7%; Score 26; DB 2; Length 134;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
   |||||
Db 121 ASAAPF 126

RESULT 11
AF2262
allophycocyanin B alpha chain [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AF2262
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigunakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2262
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-161 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA875352.1; PID:g17132786; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: apcD
C:Superfamily: phycocyanin

Query Match
Best Local Similarity 89.7%; Score 26; DB 2; Length 161;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
   |||||
Db 147 AAAAPV 152

RESULT 12
T13106
```

minor tail protein gp20 - phage N15
C:Species: phage N15
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 03-Nov-2000
C:Accession: T13106
R:Hendrix, R.W.; Ravin, V.K.; Casjens, S.R.; Ford, M.E.; Ravin, N.V.; Smirnov, I.K.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z17603
A:Accession: T13106
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-192 <HEN>
A:Cross-references: EMBL:AF064539; NID:g3192683; PID:g3192703; PIDN:AAC19056.1
C:Genetics:
A:Note: gene 20
C:Superfamily: phage lambda tail assembly protein I

Query Match 89.7%; Score 26; DB 2; Length 192;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||:|
Db 112 AAASPF 117

RESULT 13
H83587
conserved hypothetical protein PA0467 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83587
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83587
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-206 <STO>
A:Cross-references: GB:AE004484; GB:AE004091; NID:g9946320; PIDN:AAG03856.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0467

Query Match 89.7%; Score 26; DB 2; Length 206;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||:|
Db 10 AAASPF 15

RESULT 14
T36643
probable integral membrane protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36643
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21610
A:Accession: T36643
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-206 <OLI>
A:Cross-references: EMBL:AL078610; PIDN:CAB44411.1; GSPDB:GN00070; SCOEDB:SCH35.16
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCH35.16

Query Match 89.7%; Score 26; DB 2; Length 206;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||:|
Db 91 AAASPF 96

RESULT 15
C87296
beta-glucanase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: C87296
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Koehn, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: C87296
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-301 <STO>
A:Cross-references: GB:AE005673; NID:g13421537; PIDN:AAK22367.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0380

Query Match 89.7%; Score 26; DB 2; Length 301;
Best Local Similarity 83.3%; Pred. No. 3.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||:|
Db 248 ASAAPF 253

RESULT 16
C87402
conserved hypothetical protein CC1233 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: C87402
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Koehn, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: C87402
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-317 <STO>
A:Cross-references: GB:AE005673; NID:g13422561; PIDN:AAK23215.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC1233

Query Match 89.7%; Score 26; DB 2; Length 317;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||:|
Db 88 AAASPF 93

RESULT 17
T00873
hypothetical protein At2g45600 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F17K2.13
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 16-Feb-2001

C:Accession: T00873; E84892
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, March 1998
A:Description: Arabidopsis thaliana chromosome II BAC F17K2 genomic sequence.
A:Reference number: 214207
A:Accession: T00873
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-329 <ROU>
A:Cross-references: EMBL:AC003680; NID:g2979540; PID:g2979555
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Mofatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84892
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-329 <STO>
A:Cross-references: GB:AE002093; NID:g2979555; PIDN:AAC06164.1; GSPDB:GN00139
C:Genetics:
A:Gene: F17K2.13; At2g45600
A:Map position: 2

Query Match 89.7%; Score 26; DB 2; Length 329;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
||:||||
Db 82 AASAPF 87

RESULT 18
D70792
hypothetical protein Rv3691 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: D70792
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70792
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-333 <COL>
A:Cross-references: GB:AL022121; GB:AL123456; NID:g3261559; PIDN:CAA18013.1; PID:el26455
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv3691

Query Match 89.7%; Score 26; DB 2; Length 333;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
||:||||
Db 88 AASAPF 93

RESULT 19
T35939
probable transport permease protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T35939
R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, January 1999
A:Reference number: Z21551
A:Accession: T35939
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-334 <SEE>
A:Cross-references: EMBL:AL035206; PIDN:CAA22761.1; GSPDB:GN00070; SCOEDB:SC9B5.19
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC9B5.19
C:Superfamily: oligopeptide permease protein oppB

Query Match 89.7%; Score 26; DB 2; Length 334;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
||:||||
Db 39 AASAPF 44

RESULT 20
T29121
hypothetical protein SCIF2.12 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T29121
R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1998
A:Reference number: Z17215
A:Accession: T29121
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-400 <PAR>
A:Cross-references: EMBL:AL031350; NID:el1316892; PID:el1316904; PIDN:CAA20503.1
C:Genetics:
A:Note: SCIF2.12

Query Match 89.7%; Score 26; DB 2; Length 400;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
||:||||
Db 29 AASAPF 34

RESULT 21
H70931
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: H70931
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70931
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-403 <COL>
A:Cross-references: GB:AL022021; GB:AL123456; NID:g3250699; PIDN:CAA17728.1; PID:el25
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: PPE

Query Match 89.7%; Score 26; DB 2; Length 403;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||||:
Db 71 AAAAPY 76

RESULT 22

S22586

homeotic protein EVX1 - human

C:Species: Homo sapiens (man)

C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999

C:Accession: S22586

R:Paletta, A.; d'Esposito, M.; Rambaldi, M.; Acampora, D.; Balsafiore, S.; Stornaiuolo,

Nucleic Acids Res. 19, 6541-6545, 1991

A:Title: Isolation and mapping of EVX1, a human homeobox gene homologous to even-skipped

A:Reference number: S22586; MUID:92093615; PMID:1684419

A:Accession: S22586

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-407 <FAI>

A:Cross-references: EMBL:X60655; NID:g31278; PIDN:CAA43062.1; PID:g773576

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:184-240/Domain: homeobox homology <HOX>

Query Match 89.7%; Score 26; DB 2; Length 407;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||||:
Db 295 AAASPF 300

RESULT 23

A70647

probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: A70647

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: A70647

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-409 <COL>

A:Cross-references: GB:Z83867; GB:AL123456; NID:g3261695; PIDN:CAB06293.1; PID:e1299813;

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: PPE

Query Match 89.7%; Score 26; DB 2; Length 409;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||||:
Db 67 AAAAPY 72

RESULT 24

H90815

Probable kinase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: H90815

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: H90815

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-414 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA834919.1; PID:g13360960; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: ECS1496

C:Superfamily: hypothetical protein H11555

Query Match 89.7%; Score 26; DB 2; Length 414;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||||:
Db 92 AAAAPY 97

RESULT 25

D85675

Probable kinase ycfW [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: D85675

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

hler, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; DiMalanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: D85675

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-414 <STO>

A:Cross-references: GB:AE005174; NID:g12514668; PIDN:AAG55864.1; GSPDB:GN00145; UWGP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: ycfW

C:Superfamily: hypothetical protein H11555

Query Match 89.7%; Score 26; DB 2; Length 414;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||||:
Db 92 AAAAPY 97

RESULT 26

AD0645

ABC transporter integral membrane chain STY1259 [imported] - Salmonella enterica subs

C:Species: Salmonella enterica subsp. enterica serovar Typh1

A:Note: this species has also been called Salmonella typh1

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001

C:Accession: AD0645

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church

th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr

, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

A:Reference number: AB0502; PMID:11677608

A:Accession: AD0645

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-414 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD08343.1; PID:g16502388; GSPDB:GN00176

C:Genetics:

A:Gene: STY1259

C:Superfamily: hypothetical protein H11555

```
Query Match          89.7%; Score 26; DB 2; Length 414;
Best Local Similarity 83.3%; Pred. NO. 4.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
    |||||:
Db 92 AAAAPY 97

RESULT 27
C64856
probable permease ycfW - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: C64856
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: C64856
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-414 <BLAT>
A:Cross-references: GB:AE000212; GB:U00096; NID:g1787358; PIDN:AAC74202.1; PID:g1787362;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ycfW
C:Superfamily: hypothetical protein H11555
C:Keywords: transmembrane protein
F:28-44/Domain: transmembrane #status predicted <TM1>
F:271-287/Domain: transmembrane #status predicted <TM2>
F:325-341/Domain: transmembrane #status predicted <TM3>
F:378-394/Domain: transmembrane #status predicted <TM4>

Query Match          89.7%; Score 26; DB 2; Length 414;
Best Local Similarity 83.3%; Pred. NO. 4.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
    |||||:
Db 92 AAAAPY 97

RESULT 28
S12541
evx1 protein - murine sarcoma virus
C:Species: murine sarcoma virus
C:Date: 21-Nov-1993 #sequence_revision 21-Jul-1995 #text_change 24-Sep-1999
C:Accession: S12541
R:Bastian, H.; Gruss, P.
EMBO J. 9, 1839-1852, 1990
A:Title: A murine even-skipped homologue, Evx 1, is expressed during early embryogenesis
A:Reference number: S12541; MUID:90269218; PMID:1971786
A:Accession: S12541
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <BAS>
A:Cross-references: GB:X54239; NID:g50875; PIDN:CAA38145.1; PID:g50876
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:184-240/Domain: homeobox homology <HOX>

Query Match          89.7%; Score 26; DB 2; Length 416;
Best Local Similarity 83.3%; Pred. NO. 4.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
    |||||:
Db 295 AAASPF 300
```

```
RESULT 29
F84379
methylaspartate ammonia-lyase [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Mar-2001
C:Accession: F84379
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.;
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: F84379
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-418 <STO>
A:Cross-references: GB:AE004437; NID:g10581701; PIDN:AAG20402.1; GSPDB:GN00138
C:Genetics:
A:Gene: mal
C:Superfamily: Citrobacter amalonaticus methylaspartate ammonia-lyase

Query Match          89.7%; Score 26; DB 2; Length 418;
Best Local Similarity 83.3%; Pred. NO. 4.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
    |||||:
Db 262 AAAAPY 267

RESULT 30
B54843
nemo, form II - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 19-Dec-1997
R:Choi, K.W.; Benzer, S.
Cell 78, 125-136, 1994
A:Title: Rotation of photoreceptor clusters in the developing Drosophila eye requires
A:Reference number: A54843; MUID:94306509; PMID:8033204
A:Accession: B54843
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-434 <CHO>
A:Cross-references: GB:U12009
C:Genetics:
A:Gene: nmo
A:Cross-references: FLYBase:FBgn0011817
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: alternative splicing; ATP
F:38-301/Domain: protein kinase homology <KIN>
F:46-54/Region: protein kinase ATP-binding motif

Query Match          89.7%; Score 26; DB 2; Length 434;
Best Local Similarity 83.3%; Pred. NO. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
    |||||:
Db 25 AAAAPY 30

RESULT 31
E84879
probable heme A farnesyltransferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: E84879
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Mofat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
```

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84879
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-434 <STO>
A:Cross-references: GB:AE002093; NID:g6598392; PIDN:AAC27454.2; GSPDB:GN00139
C:Genetics:
A:Gene: At2g44520
A:Map position: 2

Query Match 89.7%; Score 26; DB 2; length 434;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
 |::|||
Db 418 ASAAPF 423

RESULT 32

C70780
probable PPE protein - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: C70780

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A:Authors: Sqaers, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: C70780

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-443 <COL>

A:Cross-references: GB:Z73101; GB:AL123456; NID:g3261565; PIDN:CAA97385.1; PID:g1314016
A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: PPE

C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 89.7%; Score 26; DB 2; length 443;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
 |::|||
Db 70 AAAAPY 75

RESULT 33

T48137

copper amine oxidase-like protein, incomplete - Arabidopsis thaliana

N:Alternate names: protein T4C9.110

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T48137

R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z24485

A:Accession: T48137

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-460 <BEV>

A:Cross-references: EMBL:AL080318

A:Experimental source: cultivar Columbia; BAC clone T4C9

C:Genetics:

A:Map position: 4

A:Introns: 406/3; 459/3

A:Note: T4C9.110

Query Match 89.7%; Score 26; DB 2; length 460;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
 |::|||
Db 187 AASAPF 192

RESULT 34

C70931

probable PPE protein - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: C70931

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A:Authors: Sqaers, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: C70931

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-463 <COL>

A:Cross-references: GB:AL022021; GB:AL123456; NID:g3250699; PIDN:CA17723.1; PID:e125

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: PPE

Query Match 89.7%; Score 26; DB 2; length 463;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
 |::|||
Db 66 AAAAPY 71

RESULT 35

T13425

regulatory protein K10, oocyte-specific - fruit fly (Drosophila melanogaster)

N:Alternate names: protein EG:30B8.5

C:Species: Drosophila melanogaster

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 15-Oct-1999

C:Accession: T13425; A28825

R:Murphy, L.; Harris, D.; Barrell, B.

submitted to the EMBL Data Library, April 1999

A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A:Reference number: Z17668

A:Accession: T13425

A:Status: translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-463 <MUR>

A:Cross-references: EMBL:AL009195; NID:e1355203; PID:e1202207; PIDN:CA15702.1

R:Prost, E.; Deryckere, F.; Roos, C.; Haenlin, M.; Pantesco, V.; Mohler, E.

Genes Dev. 2, 891-900, 1988

A:Title: Role of the oocyte nucleus in determination of the dorsoventral polarity of

A:Reference number: A28826

A:Accession: A28826

A:Molecule type: mRNA

A:Residues: 1-276, 'HH', 279-281, 'VDHR', 287-463 <PRO>

A:Cross-references: GB:X12836; NID:g8148; PID:g295771

C:Genetics:

A:Gene: FlyBase:fs(1)K10

A:Cross-references: FlyBase:FBgn0000810

A:Map position: X

A:Introns: 432/3

C:Keywords: DNA binding

Query Match 89.7%; Score 26; DB 2; length 463;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
|||||:
Db 53 AAAAPY 58

RESULT 36

B70932
Probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: B70932
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70932
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-468 <COL>
A:Cross-references: GB:AL022021; GB:AL123456; NID:g33250699; PIDN:CAA17730.1; PID:e125462
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: ppe

Query Match 89.7%; Score 26; DB 2; Length 468;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||||:
Db 67 AAAAPY 72

RESULT 37

T01579
heme A farnesyltransferase homolog F16B22.1 - Arabidopsis thaliana (fragment)
N:Alternate names: hypothetical protein F411.33
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 14-May-1999
C:Accession: T01579; T02407
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, July 1998
A:Description: Arabidopsis thaliana chromosome II BAC F411 genomic sequence.
A:Reference number: Z14284
A:Accession: T01579
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-469 <ROU>
A:Cross-references: EMBL:AC003672; NID:g3341671; PID:g3341672
A:Experimental source: cultivar Columbia
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul submitted to the EMBL Data Library, May 1998
A:Description: Arabidopsis thaliana chromosome II BAC F411 genomic sequence.
A:Reference number: Z14667
A:Accession: T02407
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 36-469 <ROW>
A:Cross-references: EMBL:AC004521; NID:g3128166; PID:g3128194
C:Genetics:
A:Map position: 2
A:Note: 142/2; 186/3; 232/3; 277/2; 319/2; 357/1
A:Note: F16B22.1; F411.33

Query Match 89.7%; Score 26; DB 2; Length 469;
Best Local Similarity 83.3%; Pred. No. 5.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||||:
Db 453 ASAAPF 458

RESULT 38

S08325
flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (allele BzMCC2) - maize
N:Alternate names: UDPglucose flavonoid glucosyl-transferase
C:Species: Zea mays (maize)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S08325
R:Furtek, D.; Schiefelbein, J.W.; Johnston, F.; Nelson Jr., O.E. Plant Mol. Biol. 11, 473-481, 1988
A:Title: Sequence comparisons of three wild-type Bronze-1 alleles from Zea mays.
A:Reference number: S08324
A:Accession: S08325
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-471 <FUR>
A:Cross-references: EMBL:X13501; NID:g22361; PIDN:CAA31856.1; PID:g295854
C:Genetics:
A:Gene: Bz1
A:Introns: 175/1
C:Superfamily: flavonol O3-glucosyltransferase
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 89.7%; Score 26; DB 1; Length 471;
Best Local Similarity 83.3%; Pred. No. 5.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||||:
Db 315 ASAAPF 320

RESULT 39

A54843
nemo, form I - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 24-Sep-1999
C:Accession: A54843
R:Choi, K.W.; Benzer, S. Cell 78, 125-136, 1994
A:Title: Rotation of photoreceptor clusters in the developing Drosophila eye requires
A:Reference number: A54843; MUID:94306509; PMID:8033204
A:Accession: A54843
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-477 <CHO>
A:Cross-references: GB:U12009; NID:g515669; PIDN:AAA21124.1; PID:g532558
C:Genetics:
A:Gene: nmo
A:Cross-references: FlyBase:FBgn0011817
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
F:Keywords: alternative splicing; ATP
F:38-301/Domain: protein kinase homology <KIN>
F:46-54/Region: protein kinase ATP-binding motif

Query Match 89.7%; Score 26; DB 2; Length 477;
Best Local Similarity 83.3%; Pred. No. 5.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
|||||:
Db 25 AAAAPY 30

RESULT 40

AH2761
penicillin-binding protein dac [imported] - Agrobacterium tumefaciens (strain C58, Du
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C;Accession: AH2761
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyaev, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AH2761
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-499 <KUR>
A;Cross-references: GB:AE008688; PIDN:AAL42510.1; PID:g17739929; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: dac
A;Map position: circular chromosome

Query Match 89.7%; Score 26; DB 2; Length 499;
Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
11111:
Db 314 AAAAPY 319

Search completed: December 6, 2002, 13:31:03
Job time : 16.3333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:20:29 ; Search time 8.33333 Seconds
(without alignments)
29.863 Million cell updates/sec

Title: US-10-033-526-4
Perfect score: 29
Sequence: 1 AAAAPF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|------------|--------------------|
| 1 | 29 | 100.0 | 102 | 1 | YD86_MYCTU | P71656 mycobacteri |
| 2 | 29 | 100.0 | 410 | 1 | BHB3_MOUSE | O99PV5 mus musculu |
| 3 | 29 | 100.0 | 410 | 1 | BHB3_RAT | O35779 rattus norv |
| 4 | 29 | 100.0 | 446 | 1 | TRE3_MOUSE | O64092 mus musculu |
| 5 | 29 | 100.0 | 482 | 1 | BHB3_HUMAN | O9C0J9 homo sapien |
| 6 | 29 | 100.0 | 743 | 1 | TRE3_HUMAN | P19532 homo sapien |
| 7 | 29 | 100.0 | 797 | 1 | Y900_TREPA | O83870 treponema p |
| 8 | 29 | 100.0 | 1044 | 1 | RSG1_BOVIN | P09851 bos taurus |
| 9 | 26 | 89.7 | 160 | 1 | PHAC_ANASP | P80556 anabaena sp |
| 10 | 26 | 89.7 | 272 | 1 | TNFS_CHICK | O918d8 gallus gall |
| 11 | 26 | 89.7 | 407 | 1 | EVX1_HUMAN | P49640 homo sapien |
| 12 | 26 | 89.7 | 413 | 1 | LOLE_ECOLI | P75958 escherichia |
| 13 | 26 | 89.7 | 416 | 1 | EVX1_MOUSE | P23683 mus musculu |
| 14 | 26 | 89.7 | 443 | 1 | Y878_MYCTU | Q10540 mycobacteri |
| 15 | 26 | 89.7 | 463 | 1 | R10_DROME | P13468 drosophila |
| 16 | 26 | 89.7 | 463 | 1 | Y102_MYCTU | O53951 mycobacteri |
| 17 | 26 | 89.7 | 471 | 1 | UFO2_MAIZE | P16165 zea mays (m |
| 18 | 26 | 89.7 | 515 | 1 | NU2M_BETVU | P15688 beta vulgar |
| 19 | 26 | 89.7 | 521 | 1 | YCV4_SCHPO | O74543 schizosacch |
| 20 | 26 | 89.7 | 602 | 1 | GAP1_YEAST | P19145 saccharomyc |
| 21 | 26 | 89.7 | 608 | 1 | DIP5_YEAST | P53388 saccharomyc |
| 22 | 26 | 89.7 | 647 | 1 | CN16_YEAST | P08331 escherichia |
| 23 | 26 | 89.7 | 647 | 1 | CN16_YEAST | P26265 salmonella |
| 24 | 26 | 89.7 | 652 | 1 | CN16_YEREN | P53052 yersinia en |
| 25 | 26 | 89.7 | 689 | 1 | PCEC_AERHY | Q06903 aeromonas h |
| 26 | 26 | 89.7 | 930 | 1 | PRGR_RABIT | P06186 oryctolagus |
| 27 | 26 | 89.7 | 1171 | 1 | DPOL_HSVT1 | O9YUS3 herpesvirus |
| 28 | 26 | 89.7 | 1171 | 1 | DPOL_HSVT2 | O9YUS2 herpesvirus |
| 29 | 26 | 89.7 | 1783 | 1 | RAA3_CHLRE | O9fec4 chlamydomon |
| 30 | 25 | 86.2 | 5 | 1 | SUGA_ACHDO | P19991 acheta dome |
| 31 | 25 | 86.2 | 127 | 1 | Y355_TREPA | O83374 treponema p |
| 32 | 25 | 86.2 | 141 | 1 | YEF5_YEAST | P32616 saccharomyc |
| 33 | 25 | 86.2 | 168 | 1 | RBS_SACHY | Q41373 saccharum h |

| | | | | | | |
|----|----|------|-----|---|------------|--------------------|
| 34 | 25 | 86.2 | 180 | 1 | RBS2_PEA | P00869 pisum sativ |
| 35 | 25 | 86.2 | 180 | 1 | RBS3_PEA | P07689 pisum sativ |
| 36 | 25 | 86.2 | 240 | 1 | US19_HCMVA | P09725 human cytom |
| 37 | 25 | 86.2 | 247 | 1 | TRY2_HUMAN | P07478 homo sapien |
| 38 | 25 | 86.2 | 254 | 1 | IDNO_ECOLI | P39345 escherichia |
| 39 | 25 | 86.2 | 297 | 1 | COBD_METTH | O27460 methanobact |
| 40 | 25 | 86.2 | 324 | 1 | PEP1_GADMO | P56272 gadus morhu |
| 41 | 25 | 86.2 | 326 | 1 | ATP3_IPOBA | P26360 ipomoea bat |
| 42 | 25 | 86.2 | 343 | 1 | NU2M_DINSE | O79547 dinodon sem |
| 43 | 25 | 86.2 | 346 | 1 | NU2M_PELSU | O79671 pelomedusa |
| 44 | 25 | 86.2 | 388 | 1 | HX3_XENLA | P50476 xenopus lae |
| 45 | 25 | 86.2 | 414 | 1 | CBX6_MOUSE | Q9dbys mus musculu |

ALIGNMENTS

RESULT 1
YD86_MYCTU STANDARD; PRT; 102 AA.
ID YD86_MYCTU
AC P71656;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein Rv1386 precursor.
GN Rv1386 OR MT1430 OR MTCY21B4.03.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; Pubmed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeler K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skellton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umamam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; Z80108; CAB02191.1; -.
DR EMBL; AE007014; AAK45695.1; -.
DR TIGR; MT1430; -.
DR TubercuList; Rv1386; -.
DR InterPro; IPR000084; PE_region.
DR Pfam; PF00934; PE; 1.
KW Hypothetical protein; Signal; Complete proteome.

FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 102 HYPOTHETICAL PROTEIN RV1386.
SQ SEQUENCE 102 AA; 9862 MW; 8165F09DBDB9D752 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
111111
Db 29 AAAAPF 34

RESULT 2

BHB3_MOUSE
ID BHB3_MOUSE STANDARD; PRT; 410 AA.
AC Q99PV5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Class B basic helix-loop-helix protein 3 (bHLHB3) (Differentially expressed in chondrocytes protein 2) (MDEC2).
GN BHLHB3 OR DEC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21092582; PubMed=11162494;
RA Fujimoto K., Shen M., Noshito M., Matsubara K., Shingu S., Honda K., Yoshida E., Suarida K., Matsuda Y., Kato Y.;
RT "Molecular cloning and characterization of DEC2, a new member of basic helix-loop-helix proteins.";
RL Blochem. Biophys. Res. Commun. 280:164-171(2001).
CC -!- FUNCTION: May be a transcriptional repressor that represses both basal and activated transcription.
CC -!- SUBUNIT: Homodimerize.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in skeletal muscle, brain and lung.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF TRANSCRIPTION FACTORS.
CC -!- SIMILARITY: CONTAINS 1 ORANGE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AB044090; BAB21503.1; -.
DR MGD; MGI:1930704; Bhlhb3.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS50888; HLH_2; 1.
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.
FT DNA_BIND 45 57 BASIC DOMAIN.
FT DOMAIN 58 100 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 129 175 ORANGE.
FT DOMAIN 321 373 ALA/GLY-RICH.
SQ SEQUENCE 410 AA; 43946 MW; 40A87281B08E233D CRC64;

Query Match 100.0%; Score 29; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6

Db 321 AAAAPF 326
111111

RESULT 3

BHB3_RAT
ID BHB3_RAT STANDARD; PRT; 410 AA.
AC O35779;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Class B basic helix-loop-helix protein 3 (bHLHB3) (Enhancer-of-split and hairy-related protein 1) (SHARP-1).
GN BHLHB3 OR SHARP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Cerebellum;
RX MEDLINE=98193761; PubMed=9532582;
RA Rossner M.J., Doerr J., Gass P., Schwab M.H., Nave K.-A.;
RT "SHARPs: mammalian enhancer-of-split- and hairy-related proteins coupled to neuronal stimulation.";
RL Mol. Cell. Neurosci. 10:460-475(1997).
CC -!- FUNCTION: May be a transcriptional repressor that represses both basal and activated transcription.
CC -!- SUBUNIT: Homodimerize.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Highly expressed in subregions of the brain, moderately expressed in skeletal muscle, heart. Weakly expressed in lung.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF TRANSCRIPTION FACTORS.
CC -!- SIMILARITY: CONTAINS 1 ORANGE DOMAIN.
CC -!- CAUTION: This is a conceptual translation; a frameshift was introduced in position 249 to extend the similarity with mouse ortholog.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF009329; AAB63586.1; ALT_FRAME.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS50888; HLH_2; 1.
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.
FT DNA_BIND 45 57 BASIC DOMAIN.
FT DOMAIN 58 100 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 129 175 ORANGE.
FT DOMAIN 321 344 ALA/GLY-RICH.
SQ SEQUENCE 410 AA; 43917 MW; 829705CA3A013127 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
111111
Db 321 AAAAPF 326

RESULT 4

```
TFE3_MOUSE
ID TFE3_MOUSE STANDARD: PRT: 446 AA.
AC Q64092;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Transcription factor E3 (Fragment).
GN TFE3 OR TCFE3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92123207; PubMed=1732746;
RA Roman C., Matera A.G., Cooper C., Artandi S., Blain S., Ward D.C.,
RA Calame K.;
RT "TFE3, an X-linked transcriptional activator containing basic helix-
RT loop-helix and zipper domains, utilizes the zipper to stabilize both
RT DNA binding and multimerization."
RL Mol. Cell. Biol. 12:817-827(1992).
CC -!- FUNCTION: POSITIVE-ACTING TRANSCRIPTION FACTOR THAT BINDS TO THE
CC IMMUNOGLOBULIN ENCHANCER MUE3 MOTIF. IT BINDS ALSO VERY WELL TO A
CC USE/MLE SITE. BINDING OF TFE3 TO DNA INDUCES DNA BENDING.
CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC BHLH PROTEIN.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S76673; AAB21130.1; -.
DR HSSP; P36956; 1AM9.
DR MGD; MGI:98511; Tcf3.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS50888; HLH_2; 1.
KW Transcription regulation; DNA-binding; Activator; Nuclear protein.
FT NON_TER 1
FT DOMAIN 133 144 STRONG TRANSCRIPTION ACTIVATION DOMAIN
FT 217 232 (POTENTIAL).
FT DNA_BIND 217 232 BASIC DOMAIN.
FT DOMAIN 233 273 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 282 303 LEUCINE-ZIPPER (POTENTIAL).
SQ SEQUENCE 446 AA; 47891 MW; 79115373AD7F131E CRC64;

Query Match 100.0%; Score 29; DB 1; Length 446;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 10 AAAAPF 15

RESULT 5
BHB3_HUMAN STANDARD: PRT; 482 AA.
ID BHB3_HUMAN
AC Q9C0J9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Class B basic helix-loop-helix protein 3 (BHLHB3) (Differentially
DE expressed in chondrocytes protein 2) (hDEC2) (Enhancer-of-split and
```

```
DE hairy-related protein 1) (SHARP-1).
GN BHLHB3 OR DEC2 OR SHARP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21092582; PubMed=11162494;
RA Fujimoto K., Shen M., Noshiro M., Matsubara K., Shingu S., Honda K.,
RA Yoshida E., Suardita K., Matsuda Y., Kato Y.;
RT "Molecular cloning and characterization of DEC2, a new member of basic
RT helix-loop-helix proteins."
RL Biochem. Biophys. Res. Commun. 280:164-171(2001).
RN [2]
RP FUNCTION.
RX MEDLINE=21226716; PubMed=11278948;
RA Garriga-Canut M., Roopra A., Buckley N.J.;
RT "The basic helix-loop-helix protein, SHARP-1, represses transcription
RT by a histone deacetylase-dependent and histone deacetylase-independent
RT mechanism."
RL J. Biol. Chem. 276:14821-14828(2001).
CC -!- FUNCTION: May be a transcriptional repressor that represses both
CC basal and activated transcription.
CC -!- SUBUNIT: Homodimerize.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Highly expressed in skeletal muscle and brain,
CC moderately expressed in pancreas and heart, weakly expressed in
CC placenta, lung, liver and kidney.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
CC -!- SIMILARITY: CONTAINS 1 ORANGE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB044088; BAB21502.1; -.
DR Genew; HGNC:16617; BHLHB3.
DR MIM; 606200; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS50888; HLH_2; 1.
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.
FT DNA_BIND 45 57 BASIC DOMAIN.
FT DOMAIN 58 100 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 129 175 ORANGE.
FT DOMAIN 297 431 ALA/GLY-RICH.
SQ SEQUENCE 482 AA; 50497 MW; 2BECDC2FDEB7CE14 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 482;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 380 AAAAPF 385

RESULT 6
TFE3_HUMAN STANDARD: PRT; 743 AA.
ID TFE3_HUMAN
AC P19532; Q99964; Q92757; Q92758;
DT 01-FEB-1991 (Rel. 17, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transcription factor E3.
GN TFE3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 1-219 FROM N.A., AND CHROMOSOMAL TRANSLOCATION.
RX MEDLINE=97140324; PubMed=8986805;
RA Weterman M.A.J., Wilbrink M., Geurts van Kessel A.;
RT "Fusion of the transcription factor TFE3 gene to a novel gene, PRCC,
in t(X;1)(p11;q21)-positive papillary renal cell carcinomas.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:15294-15298(1996).
RN [2]
RP SEQUENCE OF 149-743 FROM N.A., AND CHROMOSOMAL TRANSLOCATION.
RC TISSUE=Monocytes;
RX MEDLINE=97026295; PubMed=8872474;
RA Sidhar S.K., Clark J., Gill S., Hamoudi R., Crew A.J.,
RA William R., Ross M., Linehan W.M., Birdsall S., Shipley J.,
RA Cooper C.S.;
RT "The t(X;1)(p11.2;q21.2) translocation in papillary renal cell
carcinoma fuses a novel gene PRCC to the TFE3 transcription factor
gene.";
RL Hum. Mol. Genet. 5:1333-1338(1996).
RN [3]
RP REVISIONS.
RA Clark J.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 208-743 FROM N.A.
RX MEDLINE=90249724; PubMed=2338243;
RA Beckmann H., Su L.-K., Kadesch T.;
RT "TFE3: a helix-loop-helix protein that activates transcription
through the immunoglobulin enhancer mue3 motif.";
RL Genes Dev. 4:167-179(1990).
RN [5]
RP SEQUENCE OF 266-353 FROM N.A., AND CHROMOSOMAL TRANSLOCATION.
RX MEDLINE=98054131; PubMed=9393982;
RA Clark J., Lu Y.-J., Sidhar S.K., Parker C., Gill S., Smedley D.,
RA Hamoudi R., Linehan W.M., Shipley J., Cooper C.S.;
RT "Fusion of splicing factor genes PSF and Nono (p54nrb) to the TFE3
gene in papillary renal cell carcinoma.";
RL Oncogene 15:2233-2239(1997).
CC -1- FUNCTION: POSITIVE-ACTING TRANSCRIPTION FACTOR THAT BINDS TO THE
IMMUNOGLOBULIN ENHANCER MUE3 MOTIF. IT BINDS ALSO VERY WELL TO A
USE/MULT SITE. BINDING OF TFE3 TO DNA INDUCES DNA BINDING.
CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
BHLH PROTEIN.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: UBIQUITOUS IN FETAL AND ADULT TISSUES.
CC -1- DISEASE: INVOLVED IN PAPILLARY RENAL CELL CARCINOMA (PRCC) BY
CHROMOSOMAL TRANSLOCATIONS T(X;1)(P11.2;Q21.2) WHICH INVOLVES TFE3
AND PRCC; T(X;1)(P11.2;P34) WHICH INVOLVES TFE3 AND PSF, AND
INV(X)(P11.2;Q12) THAT INVOLVES TFE3 AND NONO.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X99721; CAA68061.1; -
DR EMBL: X97160; CAA65800.1; -
DR EMBL: X97161; CAA65800.1; JOINED.
DR EMBL: X97162; CAA65800.1; JOINED.
DR EMBL: X96717; CAA65478.1; -
DR EMBL: X51330; CAA35714.1; -
DR PIR: A34596; A34596.

DR PIR: S10379; S10379.
DR HSSP: P22415; 1AN4.
DR TRANSFAC: T00811; -
DR Genew: HGNC:11752; TFE3.
DR MIM: 314310; -
DR InterPro: IPR001092; HLH_basic.
DR Pfam: PF00010; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR PROSITE: PS00038; HLH_1; 1.
DR PROSITE: PS50888; HLH_2; 1.
KW Transcription regulation; DNA-binding; Activator; Nuclear protein;
KW Chromosomal translocation; Proto-oncogene.
FT DOMAIN 260 271
FT DNA_BIND 344 359
FT DOMAIN 360 400
FT DOMAIN 409 430
FT DOMAIN 575 743
FT SITE 178 179
FT SITE 295 296
FT CONFLICT 208 211
FT CONFLICT 219 219
FT CONFLICT 222 222
FT CONFLICT 443 443
FT CONFLICT 455 455
FT CONFLICT 475 475
FT CONFLICT 557 725
FT CONFLICT 593 593
FT CONFLICT 726 743
SQ SEQUENCE 743 AA; 80007 MW; 1D82B94553543949 CRC64; -
Query Match 100.0%; Score 29; DB 1; Length 743;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAADF 6
Db 137 AAADF 142
RESULT 7
Y900_TREPA STANDARD; PRT; 797 AA.
ID Y900_TREPA
AC 083870;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0900.
GN TP0900.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its


```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001259; AAC65868.1; -.
DR TIGR: TP0900; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 797 AA; 89503 MW; DEBE3440CB6D7999 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 797;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
   |||||
DB 512 AAAAPF 517

RESULT 8
RSG1_BOVIN STANDARD; PRT; 1044 AA.
ID RSG1_BOVIN
AC P09851;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ras GTPase-activating protein 1 (GTPase-activating protein) (GAP) (Ras
DE p21 protein activator) (p120GAP) (RasGAP).
GN RAS1 OR RASA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX TISSUE=Brain;
RX MEDLINE=88318957; PubMed=2842690;
RA Vogel U.S., Dixon R.A.F., Schaber M.D., Diehl R.E., Marshall M.S.,
RA Scolnick E.M., Sigal I.S., Gibbs J.B.;
RT "Cloning of bovine GAP and its interaction with oncogenic ras p21.";
RL Nature 335:90-93(1988).
CC -!- FUNCTION: INHIBITORY REGULATOR OF THE RAS-CYCLIC AMP PATHWAY.
CC -!- STIMULATES THE GTPASE OF NORMAL BUT NOT ONCOGENIC RAS P21.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 RAS-GAP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X12602; CAA31122.1; -.
DR PIR: S01966; S01966.
DR HSSP: P20936; 1WER.
DR InterPro: IPR000008; C2.
DR InterPro: IPR001849; PH.
DR InterPro: IPR001936; RasGAP.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00017; SH2; 2.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00168; C2; 1.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00616; RasGAP; 1.
DR PRINTS: PR00401; SH2DOMAIN.

```

```

DR ProDom: PD000066; SH3; 1.
DR ProDom: PD000093; SH2; 2.
DR SMART: SM00239; C2; 1.
DR SMART: SM00233; PH; 1.
DR SMART: SM00323; RasGAP; 1.
DR SMART: SM00252; SH2; 2.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS00509; RAS_GTPASE_ACTIV_1; 1.
DR PROSITE: PS50018; RAS_GTPASE_ACTIV_2; 1.
DR PROSITE: PS50001; SH2; 2.
DR PROSITE: PS50002; SH3; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
DR PROSITE: PS00499; C2_DOMAIN_1; FALSE_NEG.
DR PROSITE: PS50004; C2_DOMAIN_2; 1.
KW GTPase activation; SH2 domain; SH3 domain; Repeat.
FT DOMAIN 1 160 HYDROPHOBIC.
FT REPEAT 646 664
FT REPEAT 665 683 SH2 1.
FT DOMAIN 178 269 SH3.
FT DOMAIN 276 338 SH2 2.
FT DOMAIN 348 438
FT DOMAIN 471 574 PH.
FT DOMAIN 591 673 C2 DOMAIN.
FT DOMAIN 745 939 RAS-GAP.
FT DOMAIN 23 26 POLY-ALA.
FT DOMAIN 45 48 POLY-ALA.
FT DOMAIN 100 106 POLY-ALA.
FT DOMAIN 131 138 POLY-PRO.
FT DOMAIN 160 165 POLY-GLU.
SQ SEQUENCE 1044 AA; 115761 MW; 9B15D1223FA5A61A CRC64;

Query Match 100.0%; Score 29; DB 1; Length 1044;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
   |||||
DB 45 AAAAPF 50

RESULT 9
PHAC_ANASP STANDARD; PRT; 160 AA.
ID PHAC_ANASP
AC P80556;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Allophycocyanin alpha-B chain.
GN APCD OR AL3653.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsuno M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
RN [2]
RP SEQUENCE OF 1-28.
RX MEDLINE=96270757; PubMed=8665889;
RA Ducret A., Sidler W., Wehrli E., Frank G., Zuber H.;
RT "Isolation, characterization and electron microscopy analysis of a
RT hemidiscoidal phycobilisome type from the cyanobacterium Anabaena sp.
RT PCC 7120.";
RL Eur. J. Biochem. 236:1010-1024(1996).
CC -!- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BLUE PIGMENT-PROTEIN
CC FROM THE PHYCOBILIPROTEIN COMPLEX. ALLOPHYCOCYANIN HAS A MAXIMUM

```



```
CC      ABSORPTION AT APPROXIMATELY 654 NANOMETERS.
CC      -1- SUBUNIT: HETEROHEXAMER OF TWO ALPHA CHAINS, ONE ALPHA-B CHAIN AND
CC      THREE BETA CHAINS.
CC      -1- PTM: CONTAINS ONE COVALENTLY LINKED BILIN CHROMOPHORE.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AP003593; BAB75352.1; -
CC      DR InterPro; IPR001659; Phycobillosome.
CC      DR Pfam; PF00502; Phycobillosome; 1.
CC      DR ProDom; PD000340; Phycobillosome; 1.
CC      KW Phycobillosome; Electron transport; Photosynthesis; Bile pigment;
CC      Methylation; Complete proteome.
CC      FT INIT_MET 0 0
CC      FT MOD_RES 70 70 METHYLATION (BY SIMILARITY).
CC      FT BINDING 80 80 PHYCOCYANOBILIN CHROMOPHORE (BY
CC      FT SIMILARITY).
CC      SQ SEQUENCE 160 AA; 17680 MW; 6E814C32F2857BD6 CRC64;
CC
CC      Query Match 89.7%; Score 26; DB 1; Length 160;
CC      Best Local Similarity 83.3%; Pred. No. 77;
CC      Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
QY 1 AAAAPF 6
Db 146 AAAAPY 151
-----
RESULT 10
TNF5_CHICK STANDARD; PRT; 272 AA.
AC Q918D8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-
DE L) (CD154 protein).
GN TNFSF5 OR CD40LG OR CD40L.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn; TISSUE=Spleen;
RA Tregaskes C.A., Young J.R., Burnside J.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell
CC proliferation in the absence of co-stimulus as well as IgE
CC production in the presence of IL-4. Involved in immunoglobulin
CC class switching (By similarity).
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an
CC extracellular soluble form (By similarity).
CC -1- PTM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
```

```
DR EMBL; AJ243435; CAB95748.1; -.
DR HSSP; P29965; 1ALY.
DR InterPro; IPR003263; TNF_5.
DR InterPro; IPR003636; TNF_abc.
DR InterPro; IPR000478; TNF_family.
DR Pfam; PF00229; TNF; 1.
DR ProDom; PD002012; TNF_abc; 1.
DR ProDom; PD008600; TNF_5; 1.
DR SMART; SM00207; TNF; 1.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS50049; TNF_2; 1.
KW Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
FT CHAIN 1 272
FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT MEMBER 5, MEMBRANE FORM.
FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT MEMBER 5, SOLUBLE FORM (BY SIMILARITY).
FT CYTOPLASMIC (POTENTIAL).
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT CLEAVAGE (BY SIMILARITY).
FT POTENTIAL.
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 272 AA; 30862 MW; 5409F24A8E53CCD7 CRC64;
CC
CC      Query Match 89.7%; Score 26; DB 1; Length 272;
CC      Best Local Similarity 83.3%; Pred. No. 1.2e+02;
CC      Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
QY 1 AAAAPF 6
Db 194 AASAPF 199
-----
RESULT 11
EVX1_HUMAN STANDARD; PRT; 407 AA.
AC P49640;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox even-skipped homolog protein 1 (EVX-1).
GN EVX1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92093615; PubMed=1684419;
RA Faiella A., D'Esposito M., Rambaldi M., Acampora D.,
RA Balsafiore S., Stornaiuolo A., Mallamaci M., Miglaccio E.,
RA Gullisano M.;
RT "Isolation and mapping of EVX1, a human homeobox gene homologous to
RT even-skipped, localized at the 5' end of HOX1 locus on chromosome
RT 7."
RL Nucleic Acids Res. 19:6541-6545(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96070899; PubMed=7499236;
RA Briata P., van de Werken R., Alroidi I., Ilengo C., di Blas E.,
RA Boncinelli E., Corte G.;
RT "Transcriptional repression by the human homeobox protein EVX1 in
RT transfected mammalian cells."
RL J. Biol. Chem. 270:27695-27701(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA Bradshaw H., Hinds K., Kepler D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY PLAY A ROLE IN THE SPECIFICATION OF NEURONAL CELL
CC TYPES.
CC -----
```

```

CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- SIMILARITY: BELONGS TO THE EVEN-SKIPPED HOMEBOX FAMILY.
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60655; CAA43062.1; -.
DR EMBL; U68782; AAB07598.1; -.
DR EMBL; U68781; AAB07598.1; JOINED.
DR EMBL; AC004080; -; NOT_ANNOTATED_CDS.
DR HSSP; P14653; 1B72.
DR TRANSEAC; T02021; -.
DR Genew; HGNC:3506; EVX1.
DR MIM; 142996; -.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODOM; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Developmental protein; Homeobox; Nuclear protein.
FT DOMAIN 114 124 ASP/GLU-RICH (ACIDIC).
FT DNA_BIND 183 242 HOMEBOX.
FT DOMAIN 289 297 ALA-RICH.
FT DOMAIN 361 368 ALA-RICH.
SQ SEQUENCE 407 AA; 42439 MW; 75D696ECA1431D5A CRC64;

Query Match          89.7%; Score 26; DB 1; Length 407;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
   |||||
Db 295 AAASPF 300

RESULT 12
LOLE_ECOLI STANDARD: PRT; 413 AA.
AC P75958;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lipoprotein releasing system transmembrane protein lole.
GN LOLE OR Bll18.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=9742617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,

```

```

RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RL corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [3]
RP SEQUENCE OF 1-10, AND CHARACTERIZATION.
RX MEDLINE=20245870; PubMed=10783239;
RA Yakushi T., Masuda K., Narita S., Matsuyama S., Tokuda H.;
RT "A new ABC transporter mediating the detachment of lipid-modified
RT proteins from membranes.";
RL Nat. Cell Biol. 2:212-218(2000).
CC -I- FUNCTION: PART OF AN ATP-DEPENDENT TRANSPORT SYSTEM LOICDE
CC RESPONSIBLE FOR THE RELEASE OF LIPOPROTEINS TARGETED TO THE OUTER
CC MEMBRANE FROM THE INNER MEMBRANE. SUCH A RELEASE IS DEPENDENT OF
CC THE SORTING-SIGNAL (ABSENCE OF AN ASP AT POSITION 2 OF THE MATURE
CC LIPOPROTEIN) AND OF LOLA.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -I- SIMILARITY: BELONGS TO THE LOIC/E FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000212; AAC74202.1; -.
DR EMBL; D90747; BAA35938.1; -.
DR EcoGene; EG13441; LOLE.
DR InterPro; IPR003838; DUF214.
DR Pfam; PF02687; DUF214; 1.
KW Transport; Transmembrane; Inner membrane; Complete proteome.
FT INIT_MET 0 0
FT TRANSMEM 24 44 POTENTIAL.
FT TRANSMEM 274 294 POTENTIAL.
FT TRANSMEM 316 336 POTENTIAL.
FT TRANSMEM 376 396 POTENTIAL.
SQ SEQUENCE 413 AA; 45213 MW; C7E6A7F29D755110 CRC64;

Query Match          89.7%; Score 26; DB 1; Length 413;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
   |||||
Db 91 AAAAPY 96

RESULT 13
EVX1_MOUSE STANDARD: PRT; 416 AA.
AC P23683;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox even-skipped homolog protein 1 (EVX-1).
GN EVX1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=9026218; PubMed=1971786;
RA Bastian H., Gruss P.;
RT "A murine even-skipped homologue, Evx 1, is expressed during early
RL embryogenesis and neurogenesis in a biphasic manner.";
RL EMBO J. 9:1839-1852(1990).
RN [2]

```

```
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/c;
RX MEDLINE=92249649; PubMed=1349539;
RA Dush M.K., Martin G.R.;
RT "Analysis of mouse Evx genes: Evx-1 displays graded expression in the
RL primitive streak.";
Dev. Biol. 151:273-287(1992).
-1- FUNCTION: MAY PLAY A ROLE IN THE SPECIFICATION OF NEURONAL CELL
CC TYPES. MAY PLAY A ROLE IN THE DORSOVENTRAL SPECIFICATION OF
CC MESODERMAL CELL FATE.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DEVELOPMENTAL STAGE: SHOWS A GRADED DISTRIBUTION IN THE PRIMITIVE
CC STREAK AND IN CELLS LATERAL TO IT. IT IS NOT DETECTED IN CELLS
CC ALONG THE A-P AXIS OF THE EMBRYO ANTERIOR TO THE PRIMITIVE STREAK,
CC EXCEPT AT E7.5 WHEN THERE IS TRANSIENT EXPRESSION IN THE HEAD
CC PROXIMAL (POSTERIOR) PORTION OF THE PRIMITIVE STREAK AND CELLS
CC NEAR IT, WITH EXPRESSION LEVELS DECREASING MORE DISTALLY
CC (ANTERIORLY).
CC -1- SIMILARITY: BELONGS TO THE EVEN-SKIPPED HOMEOBOX FAMILY.
CC PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X54239; CAA38145.1; -.
DR PIR; S12541; S12541.
DR HSSP; P14653; 1B72.
DR TRANSFAC; T02023; -.
DR MGD; MGI:95461; Evx1.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEOBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEOBOX_1; 1.
DR PROSITE; PS50071; HOMEOBOX_2; 1.
KW DNA-binding; Developmental protein; Homeobox; Nuclear protein.
FT DOMAIN 114 124 ASP/GLU-RICH (ACIDIC).
FT DNA_BIND 183 242 HOMEOBOX.
FT DOMAIN 289 297 ALA-RICH.
FT DOMAIN 341 349 ALA-RICH.
FT DOMAIN 369 375 ALA-RICH.
SQ SEQUENCE 416 AA; 43198 MW; 9F0EE4F3677CD3EE CRC64;

Query Match 89.7%; Score 26; DB 1; Length 416;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAPF 6
Db 295 AAASPF 300

RESULT 14
Y878_MYCTU
ID Y878_MYCTU STANDARD; PRT; 443 AA.
AC Q10540;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical PPE-family protein Rv0878c.
GN Rv0878c OR MT0901 OR MTCY31.06C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RN Nature 393:537-544(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z73101; CAA97385.1; -.
DR EMBL; AE006977; AAK45143.1; ALT_INIT.
DR TIGR; MT0901; -.
DR Tuberculist; Rv0878c; -.
DR InterPro; IPR000030; Microbac_PPE.
DR InterPro; IPR002989; Mycobac_pentapep.
DR Pfam; PF00823; PPE; 1.
DR Pfam; PF01469; Pentapeptide_2; 4.
KW Hypothetical protein; Transmembrane; Repeat; Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 59 79 POTENTIAL.
FT TRANSMEM 181 201 POTENTIAL.
FT DOMAIN 64 73 POLY-ALA.
FT DOMAIN 81 115 ALA-RICH.
FT DOMAIN 231 270 4 X 10 AA APPROXIMATE REPEATS.
SQ SEQUENCE 443 AA; 43592 MW; C58BEC607F0675E2 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 443;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAPF 6
Db 70 AAAPY 75

RESULT 15
K10_DROME
ID K10_DROME STANDARD; PRT; 463 AA.
AC P13468; Q46075; Q9W505;
DT 01-JAN-1990 (Rel. 13, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-binding protein K10 (Female sterile protein K10).
GN FS(1)K10 OR EG:30B8.5 OR CG3218.
OS Drosophila melanogaster (Fruit fly).
```

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Prost E., Deryckere F., Roos C., Haenlin M., Pantesco V.,
RA Mohler V.,
RT "Role of the oocyte nucleus in determination of the dorsoventral
RT polarity of Drosophila as revealed by molecular analysis of the K10
RT gene.";
RL Genes Dev. 2:891-900(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Oregon-R;
RX MEDLINE=20196011; PubMed=10731137;
RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E.,
RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,
RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
RA Modolell J., Peter A., Schoettler P., Werner M., Mourikioti F.,
RA Beinert N., Dove G., Schaefer U., Jaekle H., Bucheton A.,
RA Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,
RA McMillan P.J., Salles C., Tait E.A., Valentini P., Saunders R.D.C.,
RA Glover D.M.,
RT "From sequence to chromosome: the tip of the X chromosome of D.
RT melanogaster.";
RL Science 287:2220-2222(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Venter E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: MAY HAVE A REGULATORY FUNCTION.
CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X12836; CAA31321.1; -
DR EMBL: AL009195; CAA15702.1; -
DR EMBL: AE003423; AAF45758.1; -
DR PIR: A28826; A28826.
DR FLYBase: FBgn0000810; fs(1)K10.
KW DNA-binding; Nuclear protein; Repeat.
FT DOMAIN 87 142 7 x APPROXIMATE TANDEM REPEATS.
FT REPEAT 87 94 1.
FT REPEAT 95 102 2.
FT REPEAT 103 110 3.
FT REPEAT 111 118 4.
FT REPEAT 119 126 5.
FT REPEAT 127 134 6.
FT REPEAT 135 142 7.
FT DOMAIN 284 290 POLY-PRO.
FT DNA_BIND 397 416 H-T-H MOTIF (POTENTIAL).
FT CONFLICT 277 278 PM -> HH (IN REF. 1).
FT CONFLICT 282 286 GGPPP -> VDHHR (IN REF. 1).
SQ SEQUENCE 463 AA; 51267 MW; D03C097192D1FDD0 CRC64;
Query Match 89.7%; Score 26; DB 1; Length 463;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAPF 6
Db 53 AAAAPY 58
RESULT 16
YI02_MYCTU STANDARD; PRT; 463 AA.
ID YI02_MYCTU
AC 053951;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical PPE-family protein RV1802.
GN RV1802 OR MT1851 OR MTY049.24.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tejaia F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,


```

RA Bishal W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -i- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PPE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL022021; CAA17723.1; -.
DR EMBL; AE007044; AAK46123.1; -.
DR TIGR; MT1851; -.
DR TubercuList; Rv1802; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 401 401 S -> L (IN REF. 2).
SQ SEQUENCE 463 AA; 46021 MW; EE64828BF09FA551 CRC64;

Query Match          89.7%; Score 26; DB 1; Length 463;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
   |||||
Db 66 AAAAPY 71

RESULT 17
UFO2_MAIZE
ID UFO2_MAIZE STANDARD; PRT; 471 AA.
AC P16165;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (UDP-glucose flavonoid
DE 3-O-glucosyltransferase) (Bronze-1) (Bz-MC2 allele).
GN BZ1 OR UGT71A1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Furtak D., Schiefelbein J.W., Johnston F., Nelson O.E. Jr.;
RT "Sequence comparisons of 3 wild-type bronze-1 alleles from Zea mays.";
RL Plant Mol. Biol. 11:473-481(1988).
CC -i- FUNCTION: IN THE PRESENCE OF OTHER NECESSARY COLOR FACTORS, THIS
CC GLYCOSYLATION REACTION ALLOWS THE ACCUMULATION OF ANTHOCYANIN
CC PIGMENTS.
CC -i- CATALYTIC ACTIVITY: UDP-glucose + a flavonol = UDP + flavonol
CC 3-O-D-glucoside.
CC -i- PATHWAY: Anthocyanin biosynthesis.
CC -i- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13501; CAA31856.1; -.
DR PIR; S08325; S08325.
DR MaizeDB; 13885; -.
DR InterPro; IPR002213; UDPGT.

```

```

DR Pfam; PF00201; UDPGT; 1.
DR PROSITE; PS00375; UDPGT; 1.
KW Transferase; Glycosyltransferase.
SQ SEQUENCE 471 AA; 48621 MW; 81B897410A361299 CRC64;

Query Match          89.7%; Score 26; DB 1; Length 471;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
   |||||
Db 315 ASAAPF 320

RESULT 18
NU2M_BETVU
ID NU2M_BETVU STANDARD; PRT; 515 AA.
AC P15688;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 2 (EC 1.6.5.3).
GN ND2 OR NAD2.
OS Beta vulgaris (Sugar beet).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
OX NCBI_TaxID=3555;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90318317; Pubmed-2115110;
RA Xue Y., Davies D.R., Thomas C.M.;
RT "Sugarbeet mitochondria contain an open reading frame showing
RT extensive sequence homology to the subunit 2 gene of the NADH:
RT ubiquinone reductase complex.";
RL Mol. Gen. Genet. 221:195-198(1990).
CC -i- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -i- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -i- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC -i- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-52 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16828; CAA34728.1; -.
DR EMBL; X16828; CAA34729.1; ALT_INIT.
DR PIR; S12804; DNETU2.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
KW Oxidoreductase; NAD; ubiquinone; Mitochondrion; Transmembrane.
SQ SEQUENCE 515 AA; 56274 MW; 3D60FA5FBC2820AF CRC64;

Query Match          89.7%; Score 26; DB 1; Length 515;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
   |||||
Db 100 ASAAPF 105

RESULT 19
YCV4_SCHPO
ID YCV4_SCHPO STANDARD; PRT; 521 AA.
AC Q74543;
DT 16-OCT-2001 (Rel. 40, Created)

```


DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative amino-acid permease C777.04.
 GN SPC777.04.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Useery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -i- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -i- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL031532; CAA20708.1; -
 DR InterPro: IPR002293; AA/rel_pmeasel.
 DR InterPro: IPR004840; AAC-permease.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa-permeases; 1.
 DR PROSITE: PS00218; AMINO_ACID_PERMEASE_1; 1.
 KW Hypothetical protein; transport; Amino-acid transport;
 KW Transmembrane.
 FT TRANSMEM 53 73 POTENTIAL.
 FT TRANSMEM 78 98 POTENTIAL.
 FT TRANSMEM 158 178 POTENTIAL.
 FT TRANSMEM 186 206 POTENTIAL.
 FT TRANSMEM 282 302 POTENTIAL.
 FT TRANSMEM 336 356 POTENTIAL.
 FT TRANSMEM 380 400 POTENTIAL.
 FT TRANSMEM 407 427 POTENTIAL.
 FT TRANSMEM 487 507 POTENTIAL.
 SQ SEQUENCE 521 AA; 57234 MW; ECC8E4788D97F1EF CRC64;
 Query Match 89.7%; Score 26; DB 1; Length 521;
 Best Local Similarity 83.3%; Pred. No. 2.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAPF 6
 DB 316 AAAPF 321
 RESULT 20
 GAP1_YEAST
 ID GAP1_YEAST STANDARD; PRT; 602 AA.
 AC P19145;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE General amino-acid permease GAP1.
 GN GAP1 OR YKR039W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90306009; PubMed=2194797;
 RA Jauniaux J.-C., Gresson M.;
 RT "GAP1, the general amino acid permease gene of Saccharomyces
 RT cerevisiae. Nucleotide sequence, protein similarity with the other
 RT bakers yeast amino acid permeases, and nitrogen catabolite
 RT repression.";
 RL Eur. J. Biochem. 190:39-44(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Urrastarazu L.A., Jauniaux J.-C.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -i- FUNCTION: PERMEASE FOR VARIOUS AMINO ACIDS AS WELL AS FOR GABA.
 CC -i- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -i- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X52633; CAA36858.1; -
 DR EMBL: Z28264; CAA82113.1; -
 DR PIR: S38111; S38111.
 DR SGD: S0001747; GAP1.
 DR InterPro: IPR002293; AA/rel_pmeasel.
 DR InterPro: IPR004840; AAC-permease.
 DR InterPro: IPR004841; Permease.
 DR InterPro: IPR004762; Yeast_AA_perm.
 DR Pfam: PF00324; aa-permeases; 1.
 DR TIGRfams: TIGR00913; 2A0310; 1.
 DR PROSITE: PS00218; AMINO_ACID_PERMEASE_1; 1.
 KW Transport; Amino-acid transport; Transmembrane.
 FT TRANSMEM 96 116 POTENTIAL.
 FT TRANSMEM 122 142 POTENTIAL.
 FT TRANSMEM 166 185 POTENTIAL.
 FT TRANSMEM 205 224 POTENTIAL.
 FT TRANSMEM 238 256 POTENTIAL.
 FT TRANSMEM 281 298 POTENTIAL.
 FT TRANSMEM 322 342 POTENTIAL.
 FT TRANSMEM 377 396 POTENTIAL.
 FT TRANSMEM 422 442 POTENTIAL.
 FT TRANSMEM 452 472 POTENTIAL.
 FT TRANSMEM 492 510 POTENTIAL.
 FT TRANSMEM 530 548 POTENTIAL.
 FT TRANSMEM 122 122 MISSING (IN REF. 1).
 FT TRANSMEM 189 189 S -> A (IN REF. 1).
 FT TRANSMEM 338 338 I -> V (IN REF. 1).
 FT TRANSMEM 518 518 V -> L (IN REF. 1).
 SQ SEQUENCE 602 AA; 65655 MW; 5363616447907458 CRC64;

```
Query Match          89.7%; Score 26; DB 1; Length 602;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AAAAPF 6
Db      357 AAASPF 362

RESULT 21
DIP5_YEAST STANDARD; PRT; 608 AA.
AC P53388;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dicarboxylic amino acid permease.
GN DIP5 OR YPL265M.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sigma 1278B;
RA Visser S., Grosjean S., Andre B.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansgore W.,
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Filtz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
RA Komp C., Kurt O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
RA Urestarazu L.A., Ushinsky S., Vierendeels F., Visser S., Voss H.,
RA Walsh S.V., Wambuit R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
RL Nature 387:103-105(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95802; CAA65074.1; -.
DR EMBL; Z73621; CAA98000.1; -.
DR SGD; S0006186; DIP5.
DR InterPro; IPR002293; AA/rel_pmeasel.
DR InterPro; IPR004840; AAC_permease.
DR InterPro; IPR004841; Permease.
DR InterPro; IPR004762; Yeast_AA_perm.
DR Pfam; PF00324; aa-permeases; 1.
DR TIGRfams; TIGR00913; 2A0310; 1.
DR PROSITE; PS00218; AMINO_ACID_PERMEASE_1; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 91 111
FT TRANSMEM 112 132
FT TRANSMEM 112 132
FT POTENTIAL.
```

```
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 232 252 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 323 343 POTENTIAL.
FT TRANSMEM 373 393 POTENTIAL.
FT TRANSMEM 424 444 POTENTIAL.
FT TRANSMEM 455 475 POTENTIAL.
FT TRANSMEM 502 522 POTENTIAL.
FT TRANSMEM 531 551 POTENTIAL.
FT DOMAIN 574 577 POLY-GLU.
SQ SEQUENCE 608 AA; 68097 MW; 488A382656039D0B CRC64;

Query Match          89.7%; Score 26; DB 1; Length 608;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AAAAPF 6
Db      360 AAASPF 365

RESULT 22
CN16_ECOLI STANDARD; PRT; 647 AA.
AC P08331;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 2',3'-cyclic-nucleotide 2'-phosphodiesterase precursor (EC 3.1.4.16).
GN CPDB OR B4213.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86139859; PubMed=3005231;
RA Liu J., Burns D.M., Beacham I.R.;
RT "Isolation and sequence analysis of the gene (cpdB) encoding
RT periplasmic 2',3'-cyclic phosphodiesterase.";
RL J. Bacteriol. 165:1002-1010(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
RN [3]
RP SEQUENCE OF 1-20 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=91042419; PubMed=2172762;
RA Liu J., Beacham I.R.;
RT "Transcription and regulation of the cpdB gene in Escherichia coli
RT K12 and Salmonella typhimurium LT2: evidence for modulation of
RT constitutive promoters by cyclic AMP-CRP complex.";
RL Mol. Gen. Genet. 222:161-165(1990).
RN [4]
RP SEQUENCE OF 20-31.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
CC -!- FUNCTION: THIS BIFUNCTIONAL ENZYME CATALYZES TWO CONSECUTIVE
CC REACTIONS CONVERTING 2',3'-CYCLIC-NUCLEOTIDE TO 3'-NUCLEOTIDE
CC AND THEN 3'-NUCLEOTIDE TO NUCLEIC ACID AND PHOSPHATE.
CC -!- CATALYTIC ACTIVITY: Nucleoside 2',3'-cyclic phosphate + H(2)O =
CC nucleoside 3'-phosphate.
```

```

CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: TWO KINETICALLY DISTINGUISHABLE ACTIVE SITES FOR
CC TWO CORRESPONDING SUBSTRATES CAN BE IDENTIFIED.
CC -1- SIMILARITY: BELONGS TO THE 5'-NUCLEOTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M13464; AAA23597.1; -.
DR EMBL: U14003; AAA97109.1; -.
DR EMBL: AE000492; AAC77170.1; -.
DR EMBL: X54008; CAA37954.1; -.
DR PIR: A26398; ESECPC.
DR HSSP: P07024; 20SH.
DR Ecogene: EG10160; cpdB.
DR InterPro: IPR002224; 5_nucleotidase.
DR InterPro: IPR004843; M-peptidase.
DR InterPro: IPR004844; S/T-phosphatase.
DR Pfam: PF00149; Metallophos; 1.
DR Pfam: PF02872; 5_nucleotidase; 1.
DR PROSITE: PS00785; 5_NUCLEOTIDASE_1; 1.
DR PROSITE: PS00786; 5_NUCLEOTIDASE_2; 1.
DR Hydrolyase; Multifunctional enzyme; Periplasmic; Signal;
KW Complete proteome.
FT SIGNAL 1 19
FT CHAIN 20 647
FT FT 2',3'-CYCLIC-NUCLEOTIDE 2'-
FT FT PHOSPHODIESTERASE.
FT FT CONFLICT 316 316 A -> G (IN REF. 1).
FT FT CONFLICT 528 552 GKPIDPNMFLVATNNYRAYGKFA -> ASRLIRTPCSWL
SQ SEQUENCE 647 AA; 70832 MW; 4B26DC3563473827 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 647;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAPF 6
DB 404 SAAAPF 409

RESULT 23
CN16_SALTY STANDARD; PRT; 647 AA.
ID CN16_SALTY STANDARD; PRT; 647 AA.
AC P26265;
DT 01-MAY-1992 (Rel. 22, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 2',3'-cyclic-nucleotide 2'-phosphodiesterase precursor (EC 3.1.4.16).
GN CPDB OR STM4403.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
OY 1 AAAPF 6
DB 404 SAAAPF 409

SEQUENCE FROM N.A.
STRAIN=LT2 / SGSC1412 / ATCC 700720;
MEDLINE=21534948; PubMed=11677609;
McLelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
RN [2]
RP SEQUENCE OF 1-251 FROM N.A.

```

```

RC STRAIN=LT2;
RX MEDLINE=91042419; PubMed=2172762;
RA Liu J., Beacham I.R.;
RT "Transcription and regulation of the cpdB gene in Escherichia coli
RT k12 and Salmonella typhimurium LT2: evidence for modulation of
RL constitutive promoters by cyclic AMP-CRP complex.";
Mol. Gen. Genet. 222:161-165(1990).
CC -1- FUNCTION: THIS BIFUNCTIONAL ENZYME CATALYZES TWO CONSECUTIVE
CC REACTIONS CONVERTING 2',3'-CYCLIC-NUCLEOTIDE TO 3'-NUCLEOTIDE
CC AND THEN 3'-NUCLEOTIDE TO NUCLEIC ACID AND PHOSPHATE.
CC -1- CATALYTIC ACTIVITY: Nucleoside 2',3'-cyclic phosphate + H(2)O -
CC nucleoside 3'-phosphate.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: TWO KINETICALLY DISTINGUISHABLE ACTIVE SITES FOR
CC TWO CORRESPONDING SUBSTRATES CAN BE IDENTIFIED.
CC -1- SIMILARITY: BELONGS TO THE 5'-NUCLEOTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE008906; AAL23223.1; -.
DR EMBL: X54009; CAA37956.1; -.
DR PIR: S11915; S11915.
DR StyGene: SG10065; cpdB.
DR InterPro: IPR002224; 5_nucleotidase.
DR Pfam: PF02872; 5_nucleotidase; 1.
DR PROSITE: PS00785; 5_NUCLEOTIDASE_1; 1.
DR PROSITE: PS00786; 5_NUCLEOTIDASE_2; 1.
DR Hydrolyase; Multifunctional enzyme; Periplasmic; Signal;
KW Complete proteome.
FT SIGNAL 1 19
FT CHAIN 20 647
FT FT 2',3'-CYCLIC-NUCLEOTIDE
FT FT 2'-PHOSPHODIESTERASE.
FT FT CONFLICT 51 51 G -> A (IN REF. 2).
FT FT CONFLICT 84 89 GDYMAA -> RLYGG (IN REF. 2).
FT FT CONFLICT 96 98 DVH -> GIQ (IN REF. 2).
FT FT CONFLICT 133 133 A -> G (IN REF. 2).
FT FT CONFLICT 174 174 I -> N (IN REF. 2).
SQ SEQUENCE 647 AA; 70516 MW; E9F7FE7D6681DB34 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 647;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAPF 6
DB 404 SAAAPF 409

RESULT 24
CN16_YEREN STANDARD; PRT; 652 AA.
ID CN16_YEREN STANDARD; PRT; 652 AA.
AC P53052;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 2',3'-cyclic-nucleotide 2'-phosphodiesterase precursor (EC 3.1.4.16).
GN CPDB.
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=630;
OY 1 AAAPF 6
DB 404 SAAAPF 409

SEQUENCE FROM N.A.
STRAIN=ATCC 51871 / WA-314 / Serotype O:8;
RA Truelzsch K.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS BIFUNCTIONAL ENZYME CATALYZES TWO CONSECUTIVE

```

```
CC REACTIONS CONVERTING 2',3'-CYCLIC-NUCLEOTIDE TO 3'-NUCLEOTIDE
CC AND THEN 3'-NUCLEOTIDE TO NUCLEIC ACID AND PHOSPHATE.
CC -1- CATALYTIC ACTIVITY: Nucleoside 2',3'-cyclic phosphate + H(2)O =
CC nucleoside 3'-phosphate.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO THE 5'-NUCLEOTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X85742; CA59745.1; -.
DR HSSP; P07024; 2USH.
DR InterPro; IPR002224; 5_nucleotidase.
DR InterPro; IPR004843; M-peptidase.
DR InterPro; IPR004844; S/T_phosphatse.
DR Pfam; PF00149; Metallophos; 1.
DR Pfam; PF02872; 5_nucleotidaseC; 1.
DR PROSITE; PS00785; 5_NUCLEOTIDASE_1; 1.
DR PROSITE; PS00786; 5_NUCLEOTIDASE_2; 1.
DR PROSITE; PS00786; 5_NUCLEOTIDASE_2; 1.
KW Hydrolyase; Multifunctional enzyme; Periplasmic; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 652 2',3'-CYCLIC-NUCLEOTIDE 2'-
FT PHOSPHODIESTERASE.
SQ SEQUENCE 652 AA; 71491 MW; 8781369575794E17 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 652;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 409 SAAAPF 414

RESULT 25
PPCE_AERHY STANDARD; PRT; 689 AA.
ID PPCE_AERHY STANDARD; PRT; 689 AA.
AC 006903;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Prolyl endopeptidase (EC 3.4.21.26) (Post-proline cleaving enzyme)
DE (PE).
OS Aeromonas hydrophila.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC Aeromonas.
OX NCBI_TaxID=644;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN-AK-9;
RX MEDLINE=93380910; PubMed=8370677;
RA Kanatani A., Yoshimoto T., Kitazono A., Kokubo T., Tsuru D.;
RT "Prolyl endopeptidase from Aeromonas hydrophila: cloning, sequencing,
RT and expression of the enzyme gene, and characterization of the
RT expressed enzyme.";
RL J. Biochem. 113:790-796(1993).
CC -1- FUNCTION: CLEAVES PEPTIDE BONDS ON THE C-TERMINAL SIDE OF PROLYL
CC RESIDUES WITHIN PEPTIDES THAT ARE UP TO APPROXIMATELY 30 AMINO
CC ACIDS LONG. HAS AN ABSOLUTE REQUIREMENT FOR AN X-PRO BOND IN THE
CC TRANS CONFIGURATION IMMEDIATELY PRECEDING THE PRO-Y SCISSIBLE
CC BOND.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of Pro-l-Xaa >> Ala-l-Xaa in
CC oligopeptides.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9A.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D14005; BAA03105.1; -.
DR PIR; JN0585; JN0585.
DR HSSP; P23687; 1QFM.
DR MEROPS; S09.001; -.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR004106; Peptidase_S9_N.
DR InterPro; IPR002471; Prol_endopep_ser.
DR InterPro; IPR002470; Proligo_ptase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00326; Peptidase_S9; 1.
DR Pfam; PF02897; Peptidase_S9_N; 1.
DR PRINTS; PR00862; PROLIGOPTASE.
DR PROSITE; PS00708; PRO-ENDOPEP_SER; 1.
KW Hydrolyase; Serine protease.
FT INIT_MET 0 0
FT ACT_SITE 537 537 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 656 656 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 689 AA; 76384 MW; 84DAF66E608B40EA CRC64;

Query Match 89.7%; Score 26; DB 1; Length 689;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 409 ASAAAPF 414

RESULT 26
PRGR_RABIT STANDARD; PRT; 930 AA.
ID PRGR_RABIT STANDARD; PRT; 930 AA.
AC P06186;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Progesterone receptor (PR).
DE PGR OR NR3C3.
GN Oryctolagus cuniculus (Rabbit).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87067449; PubMed=3538016;
RA Loosfelt H., Atger M., Mirrahi M., Guichon-Mantel A., Meriel C.,
RA Logeat F., Benarous R., Milgrom E.;
RT "Cloning and sequence analysis of rabbit progesterone-receptor
RT complementary DNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:9045-9049(1986).
CC -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NR3 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14547; AAA31443.1; -.
DR PIR; A25923; A25923.
```



```
DR HSSP: P06401; 1A28.
DR TRANSPAC; T00697; -.
DR InterPro; IPR000536; Hormone_rec_1lg.
DR InterPro; IPR000128; Progester_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR Pfam; PF02161; Prog_receptor; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR PRODOM; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; ZNF_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding.
FT DOMAIN 1 565 MODULATING, PRO-RICH.
FT DNA_BIND 568 633 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 568 588 C4-TYPE.
FT ZN_FING 604 628 C4-TYPE.
FT DOMAIN 678 930 STEROID-BINDING.
FT DOMAIN 184 188 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 930 AA; 98666 MW; 644FE4C13BF2F883 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 930;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 333 AAASPF 338

RESULT 27
DPOL_HSVT1 STANDARD; PRT; 1171 AA.
ID DPOL_HSVT1
AC Q9YUS3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7).
GN DPOL.
OS Herpesvirus tupaia (Strain 1) (THV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=132677;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99094630; PubMed=9880021;
RA Springfield C., Tidona C.A., Kehm R., Bahr U., Darai G.;
RT "Identification and characterization of the Tupaia herpesvirus DNA
RT polymerase gene.";
RL J. Gen. Virol. 79:3049-3053(1998).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA](N).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF074327; AAD08666.1; -.
DR InterPro; IPR002064; DNA_pol_B.
DR Pfam; PF00136; DNA_pol_B; 1.
DR Pfam; PF03104; DNA_pol_B-exo; 1.
DR PRINTS; PR00106; DNAPOLB.
DR SMART; SM00486; POLBC; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 1171 AA; 128590 MW; D2D64897FD5E70E8 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 1171;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 37 AAASAPF 42

KW DNA-binding; Nuclear protein.
SQ SEQUENCE 1171 AA; 128622 MW; CDC0480FEACFC7EC CRC64;

Query Match 89.7%; Score 26; DB 1; Length 1171;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 37 AAASAPF 42

RESULT 28
DPOL_HSVT2 STANDARD; PRT; 1171 AA.
ID DPOL_HSVT2
AC Q9YUS2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7).
GN DPOL.
OS Herpesvirus tupaia (Strain 2) (THV-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=132678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99094630; PubMed=9880021;
RA Springfield C., Tidona C.A., Kehm R., Bahr U., Darai G.;
RT "Identification and characterization of the Tupaia herpesvirus DNA
RT polymerase gene.";
RL J. Gen. Virol. 79:3049-3053(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99319892; PubMed=10392721;
RA Bahr U., Springfield C., Tidona C.A., Darai G.;
RT "Structural organization of a conserved gene cluster of Tupaia
RT herpesvirus encoding the DNA polymerase, glycoprotein B, a probable
RT processing and transport protein, and the major DNA binding protein.";
RL Virus Res. 60:123-136(1999).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA](N).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF074328; AAD08667.1; -.
DR EMBL; AF084543; AAD42936.1; -.
DR InterPro; IPR002064; DNA_pol_B.
DR Pfam; PF00136; DNA_pol_B; 1.
DR Pfam; PF03104; DNA_pol_B-exo; 1.
DR PRINTS; PR00106; DNAPOLB.
DR SMART; SM00486; POLBC; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 1171 AA; 128590 MW; D2D64897FD5E70E8 CRC64;
```



```

RESULT 29
RAA3_CHLRE
ID RAA3_CHLRE STANDARD: PRT; 1783 AA.
AC Q9FEC4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trans-splicing factor Raa3, chloroplast precursor.
GN RAA3.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN-137C / CC-125;
RX MEDLINE=21181833; PubMed=11285239;
RA Rivier C., Goldschmidt-Clermont M., Rochaix J.-D.;
RT "Identification of an RNA-protein complex involved in chloroplast
group II intron trans-splicing in Chlamydomonas reinhardtii.";
RL EMOB J. 20:1765-1773(2001).
CC -1- FUNCTION: Required for trans-splicing of exons 1 and 2 of the
chloroplast encoded psaa mRNA (a group II intron). May be required
for stability of the chloroplast RNA-protein complex in which it
is found.
CC -1- SUBUNIT: Part of a 1700 kDa complex that includes the
precursor RNA to exon 1 and the tscA RNA.
CC -1- SUBCELLULAR LOCATION: Chloroplast stroma.
CC -1- DOMAIN: The N-terminal 453 amino acids are dispensable, while the
C-terminal 630 amino acids are required for function.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF310675; AAC40000.1; -.
DR EMBL; AF310674; AAC39999.1; -.
KW Chloroplast; Transit peptide; mRNA processing; mRNA splicing.
FT TRANSIT 1 40 CHLOROPLAST (POTENTIAL).
FT CHAIN 1 40 TRANS-SPlicing FACTOR RAA3.
FT DOMAIN 112 1783 ALA-RICH.
FT DOMAIN 343 436 SER-RICH.
FT DOMAIN 478 1078 ALA-RICH.
FT DOMAIN 1310 1415 ALA-RICH.
FT DOMAIN 1416 1430 GLN-RICH.
FT DOMAIN 1496 1506 ARG-RICH.
FT DOMAIN 1678 1722 ALA-RICH.
FT DOMAIN 771 778 POLY-ALA.
FT DOMAIN 920 926 POLY-PRO.
FT DOMAIN 927 932 POLY-ALA.
FT DOMAIN 1047 1055 POLY-GLY.
FT DOMAIN 1318 1325 POLY-ALA.
FT DOMAIN 1405 1415 POLY-ALA.
FT DOMAIN 1669 1675 POLY-PRO.
SQ SEQUENCE 1783 AA; 180399 MW; 40F6206BA6EBDCDB CRC64;

Query Match 89.7%; Score 26; DB 1; Length 1783;
Best Local Similarity 83.3%; Pred. No. 7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

AC P19991;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Subesophageal ganglion pentapeptide.
OS Acheta domesticus (House cricket).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Ensifera;
OC Gryllidae; Gryllinae; Acheta.
OX NCBI_TaxID=6997;
RN [1]
RP SEQUENCE.
RA Wicker C., Wicker C.;
RT "Isolation and structure of a peptide isolated from the
subesophageal ganglion of Acheta domesticus (orthoptera).";
RL Comp. Biochem. Physiol. 88C:185-187(1987).
CC -1- SUBCELLULAR LOCATION: MAIN PEPTIDE FROM THE SUBESOPHAGEAL
CC GANGLIA.
DR PIR; JS0319; JS0319.
SQ SEQUENCE 5 AA; 476 MW; 69D76DDDDDB00000 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
Db 1 AAAPF 5

RESULT 31
Y355_TREPA
ID Y355_TREPA STANDARD: PRT; 127 AA.
AC O83374;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0355.
GN TP0355.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artlach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
spirochete.";
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001215; AAC65353.1; -.
DR TIGR; TP0355; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 127 AA; 14526 MW; 1B848EF0DAC4ADC4 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 127;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```
OY      1 AAAAPF 6
      11111
Db      83 AGAAPF 88

RESULT 32
YEF5_YEAST
ID YEF5_YEAST STANDARD; PRT; 141 AA.
AC P32616;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Hypothetical 16.5 kDa protein in GLY1-GDA1 intergenic region.
GN YEL045C OR SYGP-ORF33.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,
RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U18779; AAB64997.1; -.
DR PIR: S30832; S30832.
DR SGD: S0000771; YEL045C.
KW Hypothetical protein; ATP-binding; Transmembrane.
FT NP_BIND 15 22 ATP (POTENTIAL).
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 67 87 POTENTIAL.
SQ SEQUENCE 141 AA; 16468 MW; F6604AC5343A5D5C CRC64;

Query Match 86.2%; Score 25; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      2 AAAAPF 6
      11111
Db      38 AAAAPF 42

RESULT 33
RBS_SACHY
ID RBS_SACHY STANDARD; PRT; 168 AA.
AC Q41373;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribulose biphosphate carboxylase small chain, chloroplast precursor
DE (EC 4.1.1.39) (RUBISCO small subunit).
GN RBCS.
OS Saccharum hybrid (Sugarcane).
```

```
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Saccharum.
OX NCBI_TaxID=15819;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. H32-8560;
RX MEDLINE=93222494; PubMed=8467089;
RA Tang W., Sun S.S.;
RT "Sequence of a sugarcane ribulose-1,5-bisphosphate
RT carboxylase/oxygenase small subunit gene.";
RL plant Mol. Biol. 21:949-951(1993).
CC -i- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -i- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) = 2 3-
CC phospho-D-glycerate.
CC -i- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + O(2) =
CC 3-phospho-D-glycerate + 2-phosphoglycolate.
CC -i- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -i- SUBCELLULAR LOCATION: Chloroplast.
CC -i- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M86717; AAA33922.1; -.
DR InterPro: IPR000894; Rubisco_small.
DR Pfam: PF00101; Rubisco_small; 1.
DR PRINTS: PR00152; RUBISCO_SMALL.
DR PRODOM: PD000290; Rubisco_small; 1.
KW Photosynthesis; Carbon dioxide fixation; Photorespiration;
KW lyase; Oxidoreductase; Monooxygenase; Chloroplast; transit peptide.
FT TRANSIT 1 46 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 47 168 RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL
FT CHAIN.
SQ SEQUENCE 168 AA; 19036 MW; DED5D100B1C314F4 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 168;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      2 AAAAPF 6
      11111
Db      12 AAAAPF 16

RESULT 34
RBS2_PEA
ID RBS2_PEA STANDARD; PRT; 180 AA.
AC P00869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribulose biphosphate carboxylase small chain 3C, chloroplast
DE precursor (EC 4.1.1.39) (RUBISCO small subunit 3C) (PSS15).
DE RBCS-3C.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A. (PSS15).
RC STRAIN=cv. Progress No. 9;
```

```

RX MEDLINE=85003579; PubMed=6479146;
RA Coruzzi G., Broglie R., Edwards C., Chua N.-H.;
RT "Tissue-specific and light-regulated expression of a pea nuclear gene
RT encoding the small subunit of ribulose-1,5-bisphosphate
RT carboxylase."
RL EMBO J. 3:1671-1679(1984).
RN [2]
RP SEQUENCE FROM N.A. (RBCS-3C).
RA Fluhr R., Moses P., Morelli G., Coruzzi G., Chua N.-H.;
RT "Expression dynamics of the pea rbcS multigene family and organ
RT distribution of the transcripts."
RL EMBO J. 5:2063-2071(1986).
RN [3]
RP SEQUENCE OF 25-180 FROM N.A.
RX MEDLINE=83108917; PubMed=6296093;
RA Coruzzi G., Broglie R., Cashmore A., Chua N.-H.;
RT "Nucleotide sequences of two pea cDNA clones encoding the small
RT subunit of ribulose 1,5-bisphosphate carboxylase and the major
RT chlorophyll a/b-binding thylakoid polypeptide."
RL J. Biol. Chem. 258:1399-1402(1983).
CC -!- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -!- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) = 2 3-
CC phospho-D-glycerate.
CC -!- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + O(2) =
CC 3-phospho-D-glycerate + 2-phosphoglycolate.
CC -!- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J01257; AAA33686.1; -
DR EMBL: X00806; CAA25390.1; -
DR EMBL: X04334; CAA27865.1; -
DR PIR: A01088; RKPMS5.
DR InterPro: IPR000894; Rubisco_small.
DR Pfam: PF00101; Rubisco_small; 1.
DR PRINTS: PR00152; RUBISCO SMALL.
DR ProDom: PD000290; Rubisco_small; 1.
KW Photosynthesis; Carbon dioxide fixation; Photorespiration;
KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Transit peptide;
KW Multigene family.
FT TRANSIT 1 57 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 58 180 RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL
FT CHAIN 3C.
FT CONFLICT 25 25 A -> G (IN REF. 3).
FT CONFLICT 32 32 S -> F (IN REF. 3).
FT CONFLICT 117 117 G -> R (IN REF. 3).
SQ SEQUENCE 180 AA; 20244 MW; 212653E76E55205C CRC64;

Query Match 86.2%; Score 25; DB 1; Length 180;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAPF 6
Db 22 AAVAPF 27

RESULT 35
RBS3_PEA
ID RBS3_PEA STANDARD; PRT; 180 AA.

```

```

AC P07689; P12467;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribulose bisphosphate carboxylase small chain 3A, chloroplast
DE precursor (EC 4.1.1.39) (Rubisco small subunit 3A).
GN RBCS-3A.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Progress No. 9;
RA Fluhr R., Moses P., Morelli G., Coruzzi G., Chua N.-H.;
RT "Expression dynamics of the pea rbcS multigene family and organ
RT distribution of the transcripts."
RL EMBO J. 5:2063-2071(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87156615; PubMed=3827863;
RA Anderson S., Smith S.M.;
RT "Synthesis of the small subunit of ribulose-bisphosphate carboxylase
RT from genes cloned into plasmids containing the SP6 promoter."
RL Biochem. J. 240:709-715(1986).
RN [3]
RP SEQUENCE OF 154-180 FROM N.A.
RX MEDLINE=88296081; PubMed=3042319;
RA Hunt A.G.;
RT "Identification and characterization of cryptic polyadenylation sites
RT in the 3' region of a pea ribulose-1,5-bisphosphate carboxylase small
RT subunit gene."
RL DNA 7:329-336(1988).
CC -!- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -!- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) = 2 3-
CC phospho-D-glycerate.
CC -!- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + O(2) =
CC 3-phospho-D-glycerate + 2-phosphoglycolate.
CC -!- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X04333; CAA27864.1; -
DR EMBL: M21375; AAA33683.2; -
DR PIR: A27874; RKPMS3.
DR InterPro: IPR000894; Rubisco_small.
DR Pfam: PF00101; Rubisco_small; 1.
DR PRINTS: PR00152; RUBISCO SMALL.
DR ProDom: PD000290; Rubisco_small; 1.
KW Photosynthesis; Carbon dioxide fixation; Photorespiration;
KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Transit peptide;
KW Multigene family.
FT TRANSIT 1 57 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 58 180 RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL
FT CHAIN 3A.
SQ SEQUENCE 180 AA; 20231 MW; 33DAD53A45C0CFE7 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 180;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;

```


Query Match 86.2%; Score 25; DB 1; Length 247;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAPF 6
 1 1111
 Db 13 AVAAPF 18

RESULT 38
 IDNO_ECOLI STANDARD; PRT; 254 AA.
 AC P39345;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gluconate 5-dehydrogenase (EC 1.1.1.69) (5-keto-D-gluconate 5-reductase).
 DE IDNO OR B4266.
 GN Escherichia coli.
 OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=95334362; PubMed=7610040;
 RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L., Blattner F.R.;
 RT "Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.8 through 100 minutes."
 RL Nucleic Acids Res. 23:2105-2119(1995).
 RN [2]
 RP FUNCTION.
 RX MEDLINE=98324983; PubMed=9658018;
 RA Bausch C., Peekhaus N., Utz C., Blais T., Murray E., Lowary T., Conway T.;
 RT "Sequence analysis of the GntII (subsidiary) system for gluconate metabolism reveals a novel pathway for L-ldonic acid catabolism in Escherichia coli."
 RL J. Bacteriol. 180:3704-3710(1998).
 CC -1- FUNCTION: CATALYZES A REVERSIBLE REDUCTION OF 5-KETOGLUTAMATE TO FORM D-GLUCONATE. DEPENDENT ON NADP, ALMOST INACTIVE WITH NAD.
 CC -1- CATALYTIC ACTIVITY: D-gluconate + NAD(P)(+) = 5-dehydro-D-gluconate + NAD(P)H.
 CC -1- PATHWAY: PART OF THE GLUCONATE UTILIZATION SYSTEM GNT-II.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U14003; AAA97163.1; -.
 DR EMBL; AE000497; AAC77223.1; -.
 DR HSSP; P25529; 1AHH.
 DR Ecogene; EG12540; idno.
 DR InterPro; IPR002198; ADH_short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KW Oxidoreductase; Gluconate utilization; NADP; Complete proteome.
 FT NP_BIND 13 37 NADP (BY SIMILARITY).
 FT ACT_SITE 158 158 BY SIMILARITY.
 SQ SEQUENCE 254 AA; 27563 MW; C5AA4A044CEC1E6E CRC64;

Query Match 86.2%; Score 25; DB 1; Length 254;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAPF 6
 1 1111
 Db 60 AVAAPF 65

RESULT 39
 COBD_METTH STANDARD; PRT; 297 AA.
 ID COBD_METTH
 AC 027460;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable cobalamin biosynthesis protein cobD.
 GN COBD OR MTH1409.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J., Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K., Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D., Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R., Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S., McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M., Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum deltaH: functional analysis and comparative genomics."
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- FUNCTION: INVOLVED IN THE CONVERSION OF COBYRIC ACID TO COBINAMIDE. ADDITION OF AMINOPROPANOL ON THE F CARBOXYLIC GROUP (BY SIMILARITY).
 CC -1- PATHWAY: Cobalamin biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE COBD FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE000903; AAB85886.1; -.
 DR InterPro; IPR004485; CblB.
 DR Pfam; PF03186; CobD_CblB; 1.
 DR TIGRFAMS; TRIGR00380; cblB; 1.
 KW Cobalamin biosynthesis; Transmembrane; Complete proteome.
 FT TRANSMEM 46 66 POTENTIAL.
 FT TRANSMEM 72 92 POTENTIAL.
 FT TRANSMEM 141 161 POTENTIAL.
 FT TRANSMEM 193 213 POTENTIAL.
 FT TRANSMEM 277 297 POTENTIAL.
 SQ SEQUENCE 297 AA; 32175 MW; A37DA300FBBA4CAB CRC64;

Query Match 86.2%; Score 25; DB 1; Length 297;
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAPF 6
 1 1111
 Db 54 AVAAPF 59

RESULT 40
 PEPI_GADMO STANDARD; PRT; 324 AA.
 ID PEPI_GADMO


```

AC P56272;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pepsin IIB (EC 3.4.23.-).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
RN [1]
RP SEQUENCE, AND X-RAY CRYSTALLOGRAPHY.
RC TISSUE=Stomach;
RA Karlisen S., Hough E., Olsen R.L.;
RT "Structure and proposed amino-acid sequence of a pepsin from Atlantic
RT cod' (Gadus morhua).";
RL Acta Crystallogr. D 54:32-46(1998).
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR PDB: 1AM5; 24-DEC-97.
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam: PF00026; asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Digestion; 3D-structure.
FT ACT_SITE 32 32 BY SIMILARITY.
FT ACT_SITE 214 214 BY SIMILARITY.
FT DISULFID 45 50 BY SIMILARITY.
FT DISULFID 206 209 BY SIMILARITY.
FT DISULFID 247 280 BY SIMILARITY.
SQ SEQUENCE 324 AA; 34014 MW; EE3A6097B6941DD7 CRC64;

Query Match 86.2%; Score 25; DB 1; Length 324;
Best local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAPF 6
DB 113 AAAPF 117

```

Search completed: December 6, 2002, 13:28:35
 Job time : 9.33333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2002, 13:25:09 ; Search time 29.3333 Seconds
(without alignments)
42.146 Million cell updates/sec

Title: US-10-033-526-4
Perfect score: 29
Sequence: 1 AAAAPF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 29 | 100.0 | 128 | 4 Q96M23 | Q96M23 homo sapien |
| 2 | 29 | 100.0 | 130 | 5 Q21259 | Q21259 caenorhabdi |
| 3 | 29 | 100.0 | 235 | 2 Q936W6 | Q936W6 proionibac |
| 4 | 29 | 100.0 | 271 | 16 Q9RCZ9 | Q9RCZ9 streptomyc |
| 5 | 29 | 100.0 | 318 | 10 Q9LHU8 | Q9LHU8 oryza sativ |
| 6 | 29 | 100.0 | 348 | 8 Q9TD31 | Q9TD31 rivulus pun |
| 7 | 29 | 100.0 | 362 | 10 Q9SDG5 | Q9SDG5 oryza sativ |
| 8 | 29 | 100.0 | 389 | 16 Q92XG9 | Q92XG9 rhizobium m |
| 9 | 29 | 100.0 | 425 | 3 Q8TGV7 | Q8TGV7 tolypoclad |
| 10 | 29 | 100.0 | 482 | 4 Q8TAT1 | Q8TAT1 homo sapien |
| 11 | 29 | 100.0 | 554 | 16 Q9RS70 | Q9RS70 deinococcus |
| 12 | 29 | 100.0 | 588 | 10 Q9SDC0 | Q9SDC0 oryza sativ |
| 13 | 29 | 100.0 | 696 | 10 Q94DZ6 | Q94DZ6 oryza sativ |
| 14 | 29 | 100.0 | 722 | 12 Q8Q1N0 | Q8Q1N0 mamestra co |
| 15 | 29 | 100.0 | 723 | 5 Q26436 | Q26436 drosophila |
| 16 | 29 | 100.0 | 723 | 5 Q9W5F0 | Q9W5F0 drosophila |

| | | | | | |
|----|----|-------|------|-----------|--------------------|
| 17 | 29 | 100.0 | 723 | 5 Q9UB43 | Q9UB43 drosophila |
| 18 | 29 | 100.0 | 735 | 5 Q95TE2 | Q95TE2 drosophila |
| 19 | 29 | 100.0 | 747 | 5 Q9VWN1 | Q9VWN1 drosophila |
| 20 | 29 | 100.0 | 815 | 10 Q9LWLO | Q9LWLO oryza sativ |
| 21 | 29 | 100.0 | 821 | 12 Q86643 | Q86643 feline herp |
| 22 | 29 | 100.0 | 860 | 5 Q9NDT9 | Q9NDT9 balanus amp |
| 23 | 29 | 100.0 | 1301 | 5 Q9U122 | Q9U122 leishmania |
| 24 | 29 | 100.0 | 1598 | 5 Q95YM8 | Q95YM8 apis mellif |
| 25 | 29 | 100.0 | 3124 | 4 Q96L91 | Q96L91 homo sapien |
| 26 | 29 | 89.7 | 64 | 1 P94119 | P94119 acidianus a |
| 27 | 26 | 89.7 | 119 | 11 Q9CVT5 | Q9CVT5 mus musculu |
| 28 | 26 | 89.7 | 132 | 10 Q9AX21 | Q9AX21 oryza sativ |
| 29 | 26 | 89.7 | 134 | 17 Q9Y8T1 | Q9Y8T1 aeropyrum p |
| 30 | 26 | 89.7 | 140 | 10 Q8S7D5 | Q8S7D5 oryza sativ |
| 31 | 26 | 89.7 | 155 | 10 Q9FP74 | Q9FP74 oryza sativ |
| 32 | 26 | 89.7 | 159 | 17 Q8ZW95 | Q8ZW95 pyrobaculum |
| 33 | 26 | 89.7 | 169 | 10 Q9LG37 | Q9LG37 oryza sativ |
| 34 | 26 | 89.7 | 181 | 5 Q8W012 | Q8W012 osteriagla |
| 35 | 26 | 89.7 | 192 | 9 Q64334 | Q64334 bacterioph |
| 36 | 26 | 89.7 | 206 | 16 Q91651 | Q91651 pseudomonas |
| 37 | 26 | 89.7 | 206 | 16 Q9X8Z0 | Q9X8Z0 streptomyc |
| 38 | 26 | 89.7 | 244 | 13 Q42159 | Q42159 petromyzon |
| 39 | 26 | 89.7 | 245 | 13 Q42160 | Q42160 petromyzon |
| 40 | 26 | 89.7 | 247 | 13 Q42158 | Q42158 petromyzon |
| 41 | 26 | 89.7 | 247 | 13 Q42608 | Q42608 petromyzon |
| 42 | 26 | 89.7 | 267 | 3 Q9HEK8 | Q9HEK8 neurospora |
| 43 | 26 | 89.7 | 267 | 10 Q9LU62 | Q9LU62 arabidopsis |
| 44 | 26 | 89.7 | 271 | 16 Q9FCG6 | Q9FCG6 streptomyc |
| 45 | 26 | 89.7 | 275 | 5 Q96779 | Q96779 branchiosto |

ALIGNMENTS

RESULT 1
Q96M23 PRELIMINARY: PRT; 128 AA.
AC Q96M23; 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE CDNA FLJ31668 f1s, clone NT2R12004916.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.,
RT "NEDO human cDNA sequencing project."
RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK056230; BAB71124.1; -
SQ SEQUENCE 128 AA; 14386 MW; 7C9945A417FF0044 CRC64;

Query Match 100.0%; Score 29; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 14 AAAAPF 19

RESULT 2
Q21259 PRELIMINARY: PRT; 130 AA.

```

AC Q21259;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 14.4 kDa protein.
GN K06B9.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RL Investigating biology. The C. elegans Sequencing Consortium.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Bradshaw H., Miller N.;
RT "The sequence of C. elegans cosmid K06B9.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U50072; AAA93448.1; -.
KW Hypothetical protein.
SQ SEQUENCE 130 AA; 14364 MW; DCAB2EF12B32A93D CRC64;

Query Match          100.0%; Score 29; DB 5; Length 130;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 2 AAAAPF 7

RESULT 3
Q936W6 PRELIMINARY; PRT; 235 AA.
ID Q936W6;
AC Q936W6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cobalt transport protein Cblm.
GN Cblm.
OS Propionibacterium freudenreichii shermanii.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Propionibacterineae; Propionibacteriaceae;
OC Propionibacterium.
OX NCBI_TaxID=1752;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95189734; PubMed=7883713;
RA Satlier I., Roessner C.A., Stelowich N.J., Hardin S.H.,
RA Harris-Haller L.W., Yokubaitis N.T., Murooka Y., Hashimoto Y.,
RA Scott A.I.;
RT "Cloning, sequencing, and expression of the uroporphyrinogen III
RT methyltransferase coba gene of Propionibacterium freudenreichii
RT (shermanii).";
RL J. Bacteriol. 177:1564-1569(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Roessner C.A., Huang K., Scott A.I.;
RT "Cobalamin biosynthesis in Propionibacterium freudenreichii
RT (shermanii): Isolation and characterization of 16 B12 genes.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; U13043; AAL04403.1; -.
DR InterPro; IPR002751; Cblm.
DR Pfam; PF01891; Cblm; 1.
DR ProDom; PD005331; Cblm; 1.
DR TIGRFAMs; TIGR00123; cblm; 1.
SQ SEQUENCE 235 AA; 24534 MW; B5C9A3A4EB8A1CB6 CRC64;

Query Match          100.0%; Score 29; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 18 AAAAPF 23

RESULT 4
Q9RCZ9 PRELIMINARY; PRT; 271 AA.
ID Q9RCZ9;
AC Q9RCZ9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein SCO0912.
GN SCO0912 OR SCMI.45.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL133422; CAB62703.1; -.
KW Hypothetical protein.
SQ SEQUENCE 271 AA; 28419 MW; 068FB213A955F84D CRC64;

Query Match          100.0%; Score 29; DB 16; Length 271;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 100 AAAAPF 105

RESULT 5
Q9LH08 PRELIMINARY; PRT; 318 AA.
ID Q9LH08;
AC Q9LH08;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein.
GN Oryza sativa (Rice).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Lillipsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. NIPPONBARE;

```

```
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RL   clone:P0443E05.";
DR   Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AP001800; BAA94537.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE 318 AA; 34343 MW; AD1EF79B08225B41 CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 10; Length 318;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
    |||||
Db 221 AAAAPF 226

RESULT 6
Q9TD31 PRELIMINARY; PRT; 348 AA.
ID Q9TD31
AC Q9TD31;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE NADH dehydrogenase subunit II.
OS Rivulus punctatus.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Aplocheilidae; Rivulinae; Rivulus.
OX NCBI_TaxID=60329;
RN [1]
RP SEQUENCE FROM N.A.
RA Hirbek T., Larson A.;
RT "The evolution of diapause in the killifish family Rivulidae
RT (Atherinomorpha, Cyprinodontiformes): A molecular phylogenetic and
RT biogeographic perspective.";
RL Evolution 53:1200-1216(1999).
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
DR EMBL; AF092389; AAF02976.1; -.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 348 AA; 38539 MW; 7B6A086DC81A2F16 CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 8; Length 348;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
    |||||
Db 278 AAAAPF 283

RESULT 7
Q9SDG5 PRELIMINARY; PRT; 362 AA.
ID Q9SDG5
AC Q9SDG5;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE ESTS C19133(E10008).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0038F12.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP000836; BAA88179.1; -.
DR HSSP; P39126; IHQS.
DR InterPro; IPR001804; Isodh.
DR InterPro; IPR004434; Mito_nad_idh.
DR Pfam; PF00180; isodh; 1.
DR TIGRFAMs; TIGR00175; mito_nad_idh; 1.
DR PROSITE; PS00470; IDH_IDMH; 1.
SQ SEQUENCE 362 AA; 39442 MW; 533BE8D9DBC97860 CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 10; Length 362;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
    |||||
Db 19 AAAAPF 24

RESULT 8
Q92XG9 PRELIMINARY; PRT; 389 AA.
ID Q92XG9
AC Q92XG9;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Putative transmembrane transport protein.
GN RA1285 OR SMA2377.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymba (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RA Barlett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel M.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymba megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR EMBL; AE007314; AAK65943.1; -.
DR Transmembrane; Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 389 AA; 40252 MW; B4E899B4F4658EF3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 16; Length 389;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
    |||||
Db 354 AAAAPF 359

RESULT 9
Q8TGV7 PRELIMINARY; PRT; 425 AA.
ID Q8TGV7
AC Q8TGV7;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Serine protease protn.
OS Serine protease inflatn.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Clavicipitaceae; mitosporic Clavicipitaceae;
OC Tolypocladium.
OX NCBI_TaxID=29910;
RN [1]
```



```
RP SEQUENCE FROM N.A.
RA Plucieniczak G., Marciniak-Rusek A., Pass-Dziegielewska L.;
RT "The serine protease gene of cyclosporin-producing fungi Beauveria
RT nivea.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467982; AAL75578.1; -.
KW Protease.
SQ SEQUENCE 425 AA; 44699 MW; 8B39BD9DC5434B6A CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 3; Length 425;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 315 AAAAPF 320

RESULT 10
Q8TAT1 PRELIMINARY; PRT; 482 AA.
AC Q8TAT1;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Basic helix-loop-helix domain containing, class B, 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025968; AAH25968.1; -.
SQ SEQUENCE 482 AA; 50525 MW; 23EBD420C6B0CE03 CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 4; Length 482;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 380 AAAAPF 385

RESULT 11
Q9RS70 PRELIMINARY; PRT; 554 AA.
AC Q9RS70;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Erythromycin esterase, putative.
GN DR2257.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RA MEDLINE=20036896; Pubmed-10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Noflat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
```

```
RL Science 286:1571-1577(1999).
DR EMBL; AE002058; AAF11803.1; -.
DR TIGR; DR2257; -.
KW Complete proteome.
SQ SEQUENCE 554 AA; 58590 MW; 575A0DA75BFB31A6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 16; Length 554;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 324 AAAAPF 329

RESULT 12
Q9SDC0 PRELIMINARY; PRT; 588 AA.
AC Q9SDC0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Similar to ARABIDOPSIS THALIANA chromosome II BAC F7D8 genomic
DE sequence.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0011D01.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0667A10.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP000969; BAA8547.1; -.
DR EMBL; AP001073; BAA89568.1; -.
DR Interpro; IPR002965; P_rich_extensn.
DR PRINTS; PRO1217; PRICHEXTENS.
SQ SEQUENCE 588 AA; 62565 MW; 0EAD23FAC2ABDE0F CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 10; Length 588;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 244 AAAAPF 249

RESULT 13
Q94DZ6 PRELIMINARY; PRT; 696 AA.
AC Q94DZ6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative receptor protein kinase-like protein.
GN P0010B10.11.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
```

```
RP      SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
RL clone: P0010B10."
RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003224; BAB63567.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR01220; Lectin_legB.
DR InterPro; IPR02290; Ser_thr_pkinase.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR01559; DUFFYANTIGEN.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
DR ATP-binding; Kinase; Receptor; Transferase.
KW ATP-binding; Kinase; Receptor; Transferase.
SQ      SEQUENCE      696 AA; 74562 MW; 95BAE0C16B57FCCC CRC64;

Query Match      100.0%; Score 29; DB 10; Length 696;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAAAPF 6
        |||||
Db       7 AAAAPF 12

RESULT 14
O8QLN0 PRELIMINARY; PRT; 722 AA.
AC O8QLN0;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hoar.
OS Mamestra configurata nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxId=191492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=90/2;
RX MEDLINE=97163493; PubMed=9010313;
RA Li S., Erlandson M., Moody D., Gillott C.;
RT "A physical map of the Mamestra configurata nucleopolyhedrovirus
RT genome and sequence analysis of the polyhedrin gene.";
RL J. Gen. Virol. 78:265-271(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=90/2;
RX MEDLINE=21884635; PubMed=11886270;
RA Li Q., Donly C., Li L., Willis L.G., Theilmann D.A., Erlandson M.;
RT "Sequence and Organization of the Mamestra configurata
RT Nucleopolyhedrovirus Genome.";
RL Virology 294:106-121(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=90/2;
RA Li Q., Donly C., Li L., Willis L.G., Theilmann D.A., Erlandson M.A.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U59461; AAM09112.1; -.
SQ      SEQUENCE      722 AA; 80677 MW; 6D82DD52BBB9011F CRC64;

Query Match      100.0%; Score 29; DB 12; Length 722;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAAAPF 6
        |||||
Db       563 AAAAPF 568

RESULT 15
O26436 PRELIMINARY; PRT; 723 AA.
ID O26436;
AC O26436;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Homeobox protein VND. (Ventral nervous system defective protein).
GN VND OR NK-2 OR EG:118B3.1 OR CG6172.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95351645; PubMed=7625694;
RA Nirenberg M., Nakayama K., Nakayama N., Kim Y., Mellerick D.,
RA Wang L.H., Webber K.O., Lad R.;
RT "The NK-2 homeobox gene and the early development of the central
RT nervous system of Drosophila.";
RL Ann. N.Y. Acad. Sci. 758:224-242(1995).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; S78691; AAB34960.1; -.
DR HSSP; P22808; 1VND.
DR FlyBase; FBgn0003986; vnd.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD00010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
SQ      SEQUENCE      723 AA; 76264 MW; DC1DDE5F196B3A4D CRC64;

Query Match      100.0%; Score 29; DB 5; Length 723;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAAAPF 6
        |||||
Db       150 AAAAPF 155

RESULT 16
O9W5F0 PRELIMINARY; PRT; 723 AA.
ID O9W5F0;
AC O9W5F0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE VND protein.
GN VND OR EG:118B3.1 OR CG6172.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
```

```
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AE003418; AAF45521.1; -.
DR HSSP: P22808; 1VND.
DR FlyBase; FBgn0003986; vnd.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
KW SEQUENCE 723 AA; 76422 MW; 811F20B03A93C95E CRC64;
SQ
Query Match 100.0%; Score 29; DB 5; Length 723;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAPF 6
Db 150 AAAPF 155
RESULT 17
Q9UB43 PRELIMINARY; PRT; 723 AA.
AC Q9UB43;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE EG:118B3.1 protein.
GN VND OR EG:118B3.1 OR CG6172.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA de Pablos B., Madueno E., Modolell J.;
```

```
RT "Sequencing the distal x chromosome of Drosophila melanogaster.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Benos P.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AL031883; CAA21410.1; -.
DR HSSP: P22808; INK3.
DR FlyBase; FBgn0003986; vnd.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 723 AA; 76382 MW; 8D6932AB59470A5D CRC64;
```

```
Query Match 100.0%; Score 29; DB 5; Length 723;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAAPF 6
Db 150 AAAPF 155
```

```
RESULT 18
Q95TE2 PRELIMINARY; PRT; 735 AA.
AC Q95TE2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE LD15404P.
GN CG12527.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Xu C., Lewis S.E., Rubin G.M., Celnikier S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY059450; AAL13356.1; -.
DR FlyBase; FBgn0030977; CG12527.
SQ SEQUENCE 735 AA; 74871 MW; 8BEB8919360B7556 CRC64;
```

```
Query Match 100.0%; Score 29; DB 5; Length 735;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAAPF 6
Db 294 AAAPF 299
```

```
RESULT 19
Q9VWN1 PRELIMINARY; PRT; 747 AA.
AC Q9VWN1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CG12527 protein.
```

GN CG12527.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fostler C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003510; AAF48907.1; -.
DR FlyBase; FBgn0030977; CG12527.
SQ SEQUENCE 747 AA; 76047 MW; 3008D28A92C4EEF1 CRC64;

Query Match 100.0%; Score 29; DB 5; Length 747;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
Db 294 AAAAPF 299

RESULT 20
Q9LWLO PRELIMINARY; PRT; 815 AA.

ID Q9LWLO;
AC Q9LWLO;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Similar to Arabidopsis thaliana chromosome 1 BAC F19G10.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:p0485D09.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP001859; BAA94779.1; -.
DR InterPro: IPR003613; znf_modRING.
DR SMART; SM00504; Ubox; 1.
SQ SEQUENCE 815 AA; 85892 MW; E6F2ADBDC8E869BE CRC64;

Query Match 100.0%; Score 29; DB 10; Length 815;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
Db 517 AAAAPF 522

RESULT 21
Q86643 PRELIMINARY; PRT; 821 AA.

ID Q86643;
AC Q86643;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Glycoprotein H homolog.
OS Feline herpesvirus (Feline herpesvirus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93356636; PubMed=8394688;
RA Maeda K., Kawaguchi Y., Kamiya N., Ono M., Tohya Y., Kai C.,
RA Mikami T.;
RT "Identification and nucleotide sequence of a gene in feline
RT herpesvirus type 1 homologous to the herpes simplex virus gene
RT encoding the glycoprotein H.";
RL Arch. Virol. 132:183-191(1993).
DR EMBL; S64566; AAB27840.1; -.
DR InterPro: IPR003493; Herpes_glycopH.
DR Pfam: PF02489; Herpes_glycop_H; 1.
SQ SEQUENCE 821 AA; 92511 MW; 1F1AD4CC94404AF7 CRC64;

Query Match 100.0%; Score 29; DB 12; Length 821;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
Db 106 AAAAPF 111

RESULT 22
Q9NDT9 PRELIMINARY; PRT; 860 AA.

ID Q9NDT9;
AC Q9NDT9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE BCS-5.
GN BCS-5.
OS Balanus amphitrite (Barnacle).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Maxillopoda; Cirripedia;
OC Thoracica; Sessilia; Balanomorpha; Balanoidae; Balanus.
OX NCBI_TaxID=32267;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20314484; PubMed=10854786;

RA Okazaki Y., Shizuri Y.;
RT "Structures of six cDNAs expressed specifically at cypris larvae of
RT barnacles, Balanus amphitrite.";
RL Gene 250:127-135(2000).
DR EMBL: AB021880; BAA9547.1; -
SQ SEQUENCE 860 AA; 88965 MW; E5E905823A5905A CRC64;

Query Match 100.0%; Score 29; DB 5; Length 860;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
| | | | |
DB 319 AAAAPF 324

RESULT 23
Q9U122

ID Q9U122 PRELIMINARY; PRT; 1301 AA.

AC Q9U122;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Hypothetical 137.3 kDa protein.

GN L4768.04.

OS Leishmania major.

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI_TaxID=5664;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=FRIEDLIN;

RA Bothe G., Pohl T., Ivens A.C., Lawson D., Murphy L., Quail M.,

RA Rajandream M.A., Barrell B.G.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=FRIEDLIN;

RX MEDLINE=98146435; PubMed=9477341;

RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,

RA Smith D.F.;

RT "A physical map of the Leishmania major Friedlin genome.";

RL Genome Res. 8:135-145(1998).

DR EMBL: AL132763; CAB59861.1; -

DR InterPro: IPR001752; kinesin_motor.

DR Pfam: PF00225; kinesin_1.

DR PRINTS: PR00380; KINESINHEAVY.

DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.

KW ATP-binding; Coiled coil; Hypothetical protein; Microtubules;

KW Motor protein.

SQ SEQUENCE 1301 AA; 137312 MW; 6822477834A0B521 CRC64;

Query Match 100.0%; Score 29; DB 5; Length 1301;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
| | | | |
DB 325 AAAAPF 330

RESULT 24
Q95YM8

ID Q95YM8 PRELIMINARY; PRT; 1598 AA.

AC Q95YM8;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Mb1k-1 protein.

GN Mb1k-1.

OS Apis mellifera (Honeybee).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;

OC Apoidea; Apidae; Apis.

OX NCBI_TaxID=7460;

RN [1]

RP SEQUENCE FROM N.A.

RA Takeuchi H., Kage E., Sawata M., Kamikouchi A., Ohashi K., Ohara M.,

RA Fujiyuki T., Kunieda T., Sekimizu K., Natori S., Kubo T.;

RT "Identification of a novel gene, Mb1k-1, that encodes a putative

RT transcription factor expressed preferentially in the large-type Kenyon

RT cells of the honey bee brain.";

RL Insect Mol. Biol. 10:487-494(2001).

DR EMBL: AB047034; BAB64310.1; -

SQ SEQUENCE 1598 AA; 174929 MW; E5475BDD3ACB1EEF CRC64;

Query Match 100.0%; Score 29; DB 5; Length 1598;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
| | | | |
DB 769 AAAAPF 774

RESULT 25
Q96L91

ID Q96L91 PRELIMINARY; PRT; 3124 AA.

AC Q96L91;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE P400 SWI2/SNF2-related protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21400441; PubMed=11509179;

RA Fuchs M., Gerber J., Drapkin R., Sif S., Ikura T., Ogrzyzko V.,

RA Lane W.S., Nakatani Y., Livingston D.M.;

RT "The p400 complex is an essential E1A transformation target.";

RL Cell 106:297-307(2001).

DR EMBL: AY044869; AAK97789.1; -

DR InterPro: IPR001650; Helicase_C.

DR InterPro: IPR001005; Myb_DNA_binding.

DR InterPro: IPR000330; SNF2_N.

DR Pfam: PF00271; helicase_C; 1.

DR Pfam: PF00176; SNF2_N; 1.

DR PROSITE: PS50090; MYB_3; 1.

KW ATP-binding; Helicase.

SQ SEQUENCE 3124 AA; 340146 MW; E8F57FD6C7BD01E9 CRC64;

Query Match 100.0%; Score 29; DB 4; Length 3124;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
| | | | |
DB 1483 AAAAPF 1488

RESULT 26
P94119

ID P94119 PRELIMINARY; PRT; 64 AA.

AC P94119;

DT 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)

DE Small hydrophobic subunit of the terminal oxidase with unknown

DE homologue.

GN DOXE.

OS Acidianus ambivalens (Desulfurolobus ambivalens).

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Acidianus.

OX NCBI_TaxID=2283;


```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 3772;
RX MEDLINE=97175566; PubMed=9023221;
RA Purschke W.G., Schmidt C.L., Petersen A., Schaefer G.;
RT "The Terminal Quinol Oxidase of the Hyperthermophilic Archaeon
RT Acidianus ambivalens Exhibits a Novel Subunit Structure and Gene
RT Organization.";
RL J. Bacteriol. 179:1344-1353(1997).
DR EMBL: Y08729; CAA69982.1; -
SQ SEQUENCE 64 AA; 7196 MW; 8198E978BBB16CE9 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 64;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 30 AAAAPY 35

RESULT 27
O9CVT5 PRELIMINARY; PRT; 119 AA.
ID O9CVT5
AC O9CVT5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 1700034P14RIK protein (Fragment).
GN 1700034P14RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shlnagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boftelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK006608; BAB24672.1; -
DR MGD; MGI:1920524; 1700034P14RIK.
FT NON_TER 1
SQ SEQUENCE 119 AA; 12299 MW; 2D13BB580033671D CRC64;

Query Match 89.7%; Score 26; DB 1; Length 119;
Best Local Similarity 83.3%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 20 AAAAPY 25

RESULT 28

O9AX21 PRELIMINARY; PRT; 132 AA.
ID O9AX21
AC O9AX21;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE P0456A01.14 protein (P0435H01.2 protein).
GN P0456A01.14 OR P0435H01.2.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0456A01.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0435H01.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002899; BAB21158.1; -
DR EMBL: AP003142; BAB63509.1; -
SQ SEQUENCE 132 AA; 14544 MW; 460CE1E90B39DC0A CRC64;

Query Match 89.7%; Score 26; DB 10; Length 132;
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
Db 35 AAAAPY 40

RESULT 29
O9Y8T1 PRELIMINARY; PRT; 134 AA.
ID O9Y8T1
AC O9Y8T1;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein APE2552.
GN APE2552.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KI;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix KI.";
RL DNA Res. 6:83-101(1999).
DR EMBL: AP000064; BAA81569.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 134 AA; 13838 MW; AFD8DE992484F9B0 CRC64;

Query Match 89.7%; Score 26; DB 17; Length 134;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 AAAAPF 6
|:||||
Db 121 ASAAPF 126

RESULT 30
Q8S7D5 PRELIMINARY; PRT; 140 AA.

AC Q8S7D5; 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE Putative chloroplast chaperonin.
GN OSJNBA0057L21.3.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tsitrin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pai G.,
RA Vanaken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSJNBA0057L21 genomic sequence."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC087599; AAL79700.1; -
SQ SEQUENCE 140 AA; 14764 MW; 409F53C675EDE430 CRC64;

Query Match 89.7%; Score 26; DB 10; Length 140;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
|:||||
Db 8 AAASPF 13

RESULT 31
Q9FP74 PRELIMINARY; PRT; 155 AA.

AC Q9FP74; 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE P0458A05.15 protein (B1157F09.5 protein).
GN P0458A05.15 OR B1157F09.5.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0458A05.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
clone:B1157F09.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002870; BAB19406.1; -
SQ SEQUENCE 155 AA; 17208 MW; A3D7ED8E327629DC CRC64;

Query Match 89.7%; Score 26; DB 10; Length 155;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
|:||||
Db 77 ASAAPF 82

RESULT 32
Q8ZW95 PRELIMINARY; PRT; 159 AA.

AC Q8ZW95; 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DE P. aerophilum family 322 protein part 2, authentic frameshift.
GN PAE1905.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed-11792869;
RA Filtz-Gibson S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
aerophilum."
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009846; AAL63807.1; -
DR InterPro; IPR003675; Abi.
DR Pfam; PF02517; Abi. 1.
KW Complete proteome.

SQ SEQUENCE 159 AA; 17712 MW; 6FE1C5A0E5B844AA CRC64;

Query Match 89.7%; Score 26; DB 17; Length 159;
Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAAAPF 6
|:||||
Db 89 AAASPF 94

RESULT 33
Q9LG37 PRELIMINARY; PRT; 169 AA.

AC Q9LG37; 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE P0512G09.19 protein (P0710E05.30 protein).
GN P0512G09.19 OR P0710E05.30.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0512G09.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC

```
RT clone:P0710E05."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002836; BAB07937.1; -.
DR EMBL; AP002743; BAB16847.1; -.
SQ SEQUENCE 169 AA; 17497 MW; E3CDF68CC3D0E28A CRC64;

Query Match
Best Local Similarity 89.7%; Score 26; DB 10; Length 169;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
DE 117 SAAAPF 122

RESULT 34
O8WQ12 PRELIMINARY; PRT; 181 AA.
AC O8WQ12;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE Putative ES protein F7 (Fragment).
OS Osterlagia osterlagi.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylyda;
OC Trichostrongyloidea; Haemonchidae; Osterlagiinae; Osterlagia.
OX NCBI_TaxID=6317;
RN [1]
RP SEQUENCE FROM N.A.
RA Vercauteren I.J.R.;
RT "Identification of in vitro released excretory-secretory products of
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ427358; CAD20464.1; -.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 181 AA; 19771 MW; 28348C279C81D811 CRC64;

Query Match
Best Local Similarity 89.7%; Score 26; DB 5; Length 181;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
DE 13 ASAAPF 18

RESULT 35
O64334 PRELIMINARY; PRT; 192 AA.
AC O64334;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE GP20.
GN GENE 20.
OS Bacteriophage N15.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=40631;
RN [1]
RP SEQUENCE FROM N.A.
RA Hendrix R.W., Ravin V.K., Casjens S.R., Ford M.E., Ravin N.V.,
RA Smirnov I.K.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF064539; AAC19056.1; -.
SQ SEQUENCE 192 AA; 20149 MW; 2F1DA1B8C55B3CDE CRC64;

Query Match
Best Local Similarity 89.7%; Score 26; DB 9; Length 192;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
```

```
Db 112 AAAPF 117

RESULT 36
Q91651 PRELIMINARY; PRT; 206 AA.
AC Q91651;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein PA0467.
GN PA0467.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizner J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AE004484; AAG03856.1; -.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR004045; GST_Nterm.
DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 206 AA; 23301 MW; 42D6D51834262997 CRC64;

Query Match
Best Local Similarity 89.7%; Score 26; DB 16; Length 206;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAPF 6
DE 10 AAAPF 15

RESULT 37
Q9X820 PRELIMINARY; PRT; 206 AA.
AC Q9X820;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative integral membrane protein.
GN SC03708 OR SCH35.16.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
```

RT coellicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL078610; CAB44411.1;
SQ SEQUENCE 206 AA; 21505 MW; .62E4F58985AC2126 CRC64;

Query Match 89.7%; Score 26; DB 16; Length 206;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAPF 6
Db 91 AAASPF 96

RESULT 38
O42159 PRELIMINARY; PRT; 244 AA.
AC 042159;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Trypsinogen B1 precursor (Fragment).
GN TRYPB1.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RA Roach J.C.;
RT "The Molecular Evolution of the Vertebrate Trypsinogens.";
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF011900; AAB69656.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.128; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Signal.
FT NON_TER 1
FT SIGNAL <1 12 POTENTIAL.
FT CHAIN 13 244 TRYPSIN B1.
SQ SEQUENCE 244 AA; 25903 MW; C4582EE07E3B8007 CRC64;

Query Match 89.7%; Score 26; DB 13; Length 244;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAPF 6
Db 10 AAAAPY 15

RESULT 39
O42160 PRELIMINARY; PRT; 245 AA.
AC 042160;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Trypsinogen b2 precursor (Fragment).
GN TRYPB2.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;

RN [1]
RP SEQUENCE FROM N.A.
RA Roach J.C.;
RT "The Molecular Evolution of the Vertebrate Trypsinogens.";
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF011901; AAB69657.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.128; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Signal.
FT NON_TER 1
FT SIGNAL <1 13 POTENTIAL.
FT CHAIN 14 245 TRYPSIN B2.
SQ SEQUENCE 245 AA; 26001 MW; 9A932508B896C93E CRC64;

Query Match 89.7%; Score 26; DB 13; Length 245;
Best Local Similarity 83.3%; Pred. No. 7.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAPF 6
Db 11 AAAAPY 16

RESULT 40
O42158 PRELIMINARY; PRT; 247 AA.
AC 042158;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Trypsinogen a2 precursor.
GN TRYPB2.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RA Roach J.C.;
RT "The Molecular Evolution of the Vertebrate Trypsinogens.";
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF011898; AAB69654.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.128; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 247 TRYPSIN A2.
SQ SEQUENCE 247 AA; 26309 MW; AD73E88531970324 CRC64;

Query Match 89.7%; Score 26; DB 13; Length 247;
Best Local Similarity 83.3%; Pred. No. 7.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

• Fri Dec 6 14:09:13 2002

us-10-033-526-4.mod.rspt

Page 13

Qy 1 AAAAPF 6
| | | | |
Db 13 AAAAPY 18

Search completed: December 6, 2002, 13:30:10
Job time : 30.3333 secs

THIS PAGE BLANK (USPTO)